DOCUMENT RESUME

ED 132 773 EC 092 204

AUTHOR Miller, Susan
TITLE Tests Used with Exceptional Children: Annotated Bibliography.
INSTITUTION Drake Univ., Des Moines, Iowa.
PUB DATE 31 Jul 75
NOTE 98p.

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS *Annotated Bibliographies; Diagnostic Tests; Early Childhood Education; Elementary Secondary Education; Exceptional Child Education; *Handicapped Children; *Measurement Instruments; Screening Tests; *Standardized Tests; Student Evaluation

ABSTRACT

Provided is an annotated bibliography of 86 measurement instruments which are used in screening and formulating diagnoses for exceptional children. The following components are included in each test summary: title, date of the most recent revision, author, range (in terms of chronological age or grade placement), administration (group or individual, time required, and training of administrator), brief description, development (including standardization), reliability (including validity), and name and address of distributor or publisher. Tests are categorized into the following areas: visual and auditory acuity, intelligence, social-emotional, early screening, speech, language and concepts, auditory perception, visual perception and visual-motor integration, gross motor, learning disabilities, reading readiness, reading, arithmetic, and general achievement. (IM)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED132773

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Tests Used with Exceptional Children
Annotated Bibliography

Susan Miller
State Educational Media Consultant
July 31, 1975

EC097901

The purpose of this paper is to provide an alphabetized annotated bibliography of some measurement instruments which are used in screening and formulating diagnoses. Descriptive information is provided for each test. Information about these tests was obtained from test copies, manuals, and publisher catalogues. Since tests are frequently revised and new forms or manuals issued, anyone interested in using a particular test should consult the publisher's most current catalogue before purchase.

In some cases, evaluative statements from various sources were included in the summary of the test. Generally, these evaluations have been taken from one of the Mental Measurements Yearbooks, O.K. Buros, Editor. These and other sources of evaluative reviews are listed in the references. It is inevitable in a condensation of this type that important information may be omitted. The interested educator is urged to consult the original reference, such as the test manual or the review, before making any final evaluations. A few frequently used tests have been included with reviewer's comments to point out the inadvisability of using those particular tests.

Inclusion of any test in this bibliography should not be construed as a recommendation by the Missouri Department of Elementary and Secondary Education.

Several components are included in each test summary. The following points should be considered when reading the summaries:

Title. The title of the test is not always the name most frequently used when educators refer to the test. The author's name is often used in place of the actual title, e.g. the Developmental Test of Visual Motor Integration is often called the "Beery."

Date. The date refers to the most recent revision of a test when this information was available. The original edition of a test may have been published in 1933, and a revised edition made available in 1972. In some cases, the original and revised versions may be virtually identical; however, the inference is made that the "revised" edition constitutes an update of the test. The interested educator should refer to both editions and critics' reviews to see what changes, if any, were made in the original edition.

Author. In cases where more than one person authored a test, all names listed in the source are included.

Range. The range of the test may be stated in terms of chronological age (ages 3-6) or grade placement (grades K-12). Information on levels available is included after the age or grade range. For example, "Grades 1-6; 2 levels" indicates that the test is available for two different grade ranges. One level may be for grades 1-3 and a second level for grades 4-6. The catalogue should be consulted for more information on the range of each level.

Administration. A test will be described as a group administered test or an individually administered test. Of course, a group test could also be administered to an individual student. Following this information is the estimated time for test administration. If alternate forms are available which are intended to be equivalent in content, difficulty, etc., this is stated after the administration time. If a trained examiner is required to administer the test, this is mentioned after time or forms.

Description. In this section tests may be referred to by their initials only. For example, Metropolitan Readiness Tests may be referred to as MRT. A general description of the test, as well as reviewers' comments, are provided in this section. A comment from a review in the Mental Measurements Yearbook includes the critic's last name, the edition number and MMY, e.g. (Smith, 7th MMY). Other reviews are referred to by the author's name and the year of publication, e.g. (Jones, 1970). A few tests are described with no further information provided. These tests are well known but would not be used often in a public school settings as few personnel would be trained to administer and interpret them.

Development. Information about the population upon which the test was standardized is included under "Development." In many instances, the number of individuals in the sample may be approximate. Information on the normative population has been condensed for tests when extensive descriptions were provided in the manual. If a test is being considered for use, the manual or technical report should be consulted. It should be determined if the sample upon which the test was standardized is representative (in terms of race, community size, socioeconomic class, etc.) of the students to be assessed with that test.

Reliability. It is important that information on reliability in the test manual or technical report be evaluated if one is considering a particular test. It would be impossible to meaningfully condense numerous studies related to reliability which are available for some tests. Other tests do not report information on reliability. It may require many years before adequate reliability information is accumlated. Lack of such information should not automatically result in discontinued use of the test. However, careful consideration should be given before using such a test to make important decisions regarding a student's educational future.

Validity. The comments under "Reliability" are also appropriate for "Validity."

Publisher. This section includes the most recent available name and address of the publisher or distributor of the test.

Visual and Auditory Acuity

Vision and hearing screening tests may be given by someone trained to use these specialized instruments. This individual would most likely be the school nurse. Those children who show possible defects in vision or hearing on the screening test should be referred to the appropriate specialist for a thorough and intensive professional examination.

The following are devices which may be used for screening visual acuity:

- AO School Vision Screening Test
- Keystone-Telebinocular
- Ortho-Rater
- Titmus Vision Tester

The following companies manufacturer audiometers which are used for screening auditory acuity:

- Ambco Electronics
- Beltone Electronics Corporation
- Eckstein Brothers, Inc.
- Grason-Stadler
- Maico Hearing Instruments
- Tracor, Inc.
- Zenith Hearing Instrument Corporation

Intelligence

The tests listed in this category are usually administered to obtain an overall estimate of the subject's general level of functioning capacity. They are not designed for the purpose of analyzing the characteristics of the subject's learning problems or for determining appropriate remedial strategies.

- California Short-Form Test of Mental Maturity
- Cognitive Abilities Test (Primary Battery)
- Hiskey-Nebraska Test of Learning Aptitude
- Lorge-Thorndike Intelligence Tests, Multi-Level Edition
- Lorge-Thorndike Intelligence Test, Primary Battery
- Otis-Lennon Mental Ability Test
- Peabody Picture Vocabulary Test
- Slosson Intelligence Test
- SRA Primary Mental Abilities Test

Stanford-Binet Intelligence Scale, Third Revision
Wechsler Intelligence Scale for Children-Revised
Wechsler Preschool and Primary Scale of Intelligence

Social-Emotional

Although social-emotional problems are often best assessed through observation and other informal techniques, there are formal instruments available to assess the appropriateness of certain behaviors. Most of the tests in this section can be administered in a nonclinical setting.

Adaptive Behavior Scales
Bender-Gestalt Test
Cain-Levine Social Competency Scale
Devereux Adolescent Behavior Rating Scale
Devereux Child Behavior Rating Scale
Devereux Elementary School Behavior Rating Scale
Vineland Social Maturity Scale
What I Like to Do

Early Screening

According to the handbook develoepd by the Missouri Department of Elementary and Secondary Education, "Guidelines for an Early Childhood Screening Program," the definition of screening is the "use of relatively simple devices administered on initial contact with the population which are valid and reliable in terms of determing relative normalcy." The instruments in this section are designed to assist in identifying the needs of children at an early age - usually prior to entering kindergarten.

Cooperative Preschool Inventory
Denver Developmental Screening Test
First Grade Screening Test
Pre-School Kindergarten Readiness Inventory
Screening Test of Academic Readiness
Vane Kindergarten Test

Speech

Often a distinction is not made between "speech" and "language." Speech disorders are generally considered to be difficulties in sound production. These would include articulatory defects, stuttering, cleft-palate speech, or vocal defects (faulty pitch, quality or intensity).

Arizona Articulation Proficiency Scale
Ohio Tests of Articulation and Perception of Sounds
Predictive Screening Test of Articulation
Riley Articulation and Language Test
Templin-Darley Tests of Articulation

Language and Concepts

Language is the system used for human communication. Unlike speech, which is more related to mechanical reproduction of sounds, language is related to the expression of thoughts and understanding of ideas. The development of language and concept formation is dependent upon a background of experience.

Basic Concept Inventory, Field Research Edition
Boehm Test of Basic Concepts
Illinois Test of Psycholinguistic Abilities
Northwestern Syntax Screening Test
Preschool Language Scale
Utah Test of Language Development

Auditory Perception

The tests included here are probably best used for assessing highly specialized auditory skills such as auditory discrimination or auditory memory. They should be given after auditory acuity has been checked. The use of these various auditory tests to identify children as potential school failures, as readiness for school measures, or as predictors of reading achievement is not recommended.

Auditory Discrimination Test
Goldman-Fristoe-Woodcock Auditory Skills Test Battery
Goldman-Fristoe-Woodcock Test of Auditory Discrimination
Kindergarten Auditory Screening Test
Ohio Tests of Articulation and Perception of Sounds

Visual Perception and Visual-Motor Integration

Like the auditory perception tests, visual perception tests are most efficiently used to assess specific visual skills such as visual discrimination or visual memory. They should be administered after visual acuity has been checked. It also should not be assumed that these tests may be used efficiently as predictors of reading achievement or overall school success. In selecting and using one of these tests, one whould take into consideration the fact that most tests of visual perception also involve fine-motor skills; thus, the term "visual-motor integration."

Bender Gestalt Tests
Benton Revised Visual Retention Test
Developmental Test of Visual-Motor Integration
Marianne Frostig Developmental Test of Visual Perception
Motor-Free Visual Perception Test
Perceptual Forms Test
Primary Visual Motor Test
Purdue Perceptual Motor Survey

Gross Motor

Most assessment of gross motor skill employs informal observation techniques. Gross motor activities are those activities which primarily involve the movement of large muscles.

Lincoln-Oseretsky Motor Development Scale
Purdue Perceptional Motor Survey

Learning Disabilities

The tests listed in the "Learning Disabilities" category are designed to measure a variety of abilities which may interfere with success in school. These may include a combination of skills such as visual and auditory perception, gross and fine motor or language skills which may be assessed by other devices also.

- Behavior Rating Scale
- Meeting Street School Screening Test
- Pre-Reading Screening Procedures
- Pupil Rating Scale
- Screening Tests for Identifying Children with Specific Language Disability
- Specific Language Disability Test
- Valett Developmental Survey of Basic Learning Abilities

Reading Readiness

Readiness tests are group tests and are commonly given at the end of kindergarten or beginnning of first grade. They have been developed to evaluate the child's readiness for academic tasks, especially reading.

- Gates-MacGinitie Reading Tests: Readiness Skills
- Lee-Clark Reading Readiness Test
- Metropolitan Readiness Tests
- Murphy-Durrell Reading Readiness Analysis

Reading

Tests used to assess reading performance may be survey tests or diagnostic tests. A survey reading test is designed to give an indication of the level of difficulty at which the subject can read or a general level of comprehension. A diagnostic test is designed to give a more detailed analysis of reading skills. General achievement tests also may include reading subtests which assess overall level of reading ability.

Survey

- Classroom Reading Inventory
- Durrell Listening-Reading Series
- Gates-MacGinitie Reading Tests
- New Developmental Reading Tests (Intermediate)
- New Developmental Reading Tests (Primary)
- Standard Reading Inventory

Diagnostic

- Botel Reading Inventory
- Diagnostic Reading Scales
- Doren Diagnostic Reading Test
- Durrell Analysis of Reading Difficulty

Gates-McKillop Reading Diagnostic Test
Gilmore Oral Reading Test
Gray Oral Reading Test
McCullough Word-Analysis Tests
Silent Reading Diagnostic Tests
Standord Diagnostic Reading Test
Woodcock Reading Mastery Tests

Arithmetic

Standardized tests of arithmetic generally fall into two types. One type of test assesses a subject's specific strengths and weakensses. Some of the tests of this type provide a profile of skills which may be helpful in planning remedial programs. The other type of standardized arithmetic test is the general achievement test which includes an arithmetic subtest as part of the complete battery. This type of assessment tests overall level of arithmetic functioning but does not diagnose specific skills or deficits.

Diagnostic Chart for Fundamental Processes in Arithmetic
Diagnostic Tests and Self-Helps in Arithmetic
KeyMath Diagnostic Arithmetic Test
Stanford Diagnostic Arithmetic Test

General Achievement

Tests usually contain arithmetic and reading subtests and may include a spelling subtest. General achievement tests survey overall level of functioning and thus are not most efficient to use for diagnostic purposes.

Peabody Individual Achievement Test
Stanford Early School Achievement (Level I)
Wide Range Achievement Test

Title: Adaptive Behavior Scales

Date: 1969

Authors: Kazuo Nihira, Ray Foster, Max Shellhaas and Henry Leland

Range: Ages 3 and over; 2 levels

Administration: Individual rating; 20-25 minutes

Description: The ABS is a set of 111 items covering 24 areas of social and personal behavior. Its major purpose is to evaluate the subject's effectiveness in coping with environmental demands. The ABS was designed to follow the definition of adaptive behavior of the American Association on Mental Deficiency. It was developed for use with emotionally disturbed and mentally retarded subjects. The rating scale may be completed by any individual who is well acquainted with the subject. Miller (7th MMY) stated that the ABS should be useful in providing essential information for making decisions about the mentally retarded. According to Semmell (7th MMY), the ABS is potentially useful but needs more research on reliability and validity, and more extensive standardization.

Development: The ABS was standardized on a sample of 2,800 patients from 63 residential institutions for the retarded in the U.S. The sample was stratified by sex, intelligence, and age.

Reliability: Inter-rater reliabilities reportedly ranged from .40 to .86 for the adult level.

Validity: According to Miller (7th MMY), some studies are cited as evidence of validity and future studies will probably continue to validate the ABS.

Publisher:
American Association on Mental Deficiency
5201 Connecticut Avenue N.W.
Washington, D.C. 20015

Title: Arizona Articulation Proficiency Scale

Date: 1970

Author: Janet Barker Fudala

Range: Mental ages 2-14 and over

Administration: Individual; 20 minutes

Description: The AAPS is a numerical scale of articulatory proficiency. It is based on the assumption that the more frequently a sound is misarticulated, the greater is the articulatory problem.

Development: The standardization population consisted of 700 boys and girls, ages 3-0 to 11-11, in the Seattle Public Schools. Children with gross deviations in hearing, mental ability, emotional stability, or neurological functioning were eliminated. Children from 16 preschools were also included. No clear description is given of the sample.

Reliability: A test-retest reliability coefficient of .96 with a one-week test interval was reported for 105 children.

Validity: A correlation of .92 was reported between the AAPS and 10 speech pathologists' ratings of 45 children.

Publisher:
Western Psychological Services
12031 Wilshire Boulevard
Los Angeles, California 90025

Title: Auditory Discrimination Test

Date: 1973

Author: Joseph M. Wepman

Range: Ages 5-8 (and above)

Administration: Individual; 5 minutes, 2 forms

Description: The Wepman ADT is used for measuring sound discrimination among young children and has also been used in diagnosing older children experiencing speech or reading difficulties. Various studies have indicated that children with articulation problems are very likely to score low in this test. DiCarlo (6th MMY) reviewed the ADT, giving it a good overall rating.

Development: The standardization population consisted of 1,000 children ages 5 to 8, from New York, Illinois and Iowa. Socioeconomic status, ethnic group, urban-rural classification, parental education level and other variables were considered.

Reliability: Test-retest reliability coefficient of .91 and .95 were reported for samples of 109 and 279. An equivalent-form coefficient of .92 was reported. A coefficient of .62 was reported on the difficulty of the phoneme comparisons for a sample of 67.

Validity: Various studies have been undertaken to investigate the validity. These were summarized in the ADT manual.

Publisher: Language Research Associates, Inc.
175 East Delaware Place
Chicago, Illinois 60611

Title: Basic Concept Inventory, Field Research Edition

Date: 1967

Author: Siegfried E. Engelmann

Range: Preschool and Kindergarten

Administration: Individual; 15 minutes

Description: The BCI is a criterion-referenced checklist of basic concepts frequently used in verbal directions which are considered necessary for a child preparing for beginning academic tasks. It was primarily designed for use with "experientially-deprived and emotionally-disturbed children." Its use with slow learners and the mentally retarded is also mentioned. Much of the manual is devoted to suggestions for instruction to the classroom teacher. Caution must be used since the BCI is a research edition. However, McCandless (7th MMY) felt that the BCI would be useful to any teacher interested in diagnostic testing. McCarthy (7th MMY) believed that the BCI has some merit if used cautiously.

Development: No information was available.

Reliability: No information was available.

Validity: No information was available.

Publisher: Follett Publishing Company
1010 West Washington Boulevard
Chicago, Illinois 60607

Title: Behavior Rating Scale

Date: 1975

Authors: Veralee B. Hardin and Robert F. Busch

Range: Primary grades

Administration: Teacher rating; 10-15 minutes

Description: The BRS is a checklist to be used by the classroom teacher identifying children with possible learning disabilities. The BRS checklist consists of 60 observable characteristics of the primary grade student. The four areas included are Visual Perception, Auditory Perception, Language and Motor Development.

Development: As this is a new instrument, information is not yet available on development and reliability.

Validity: One study has been done at this time which indicated that the BRS, if used with the Cognitive Abilities Test (group intelligence test) and Knowledge of Letters Names and Sounds (from the Stanford Early School Achievement Test), had a correlation coefficient of .59.

Publisher: Lucas Brothers Publishers
909 Lowry
Columbia, Missouri 65201

Title: Bender-Gestalt Test

Date: 1946

Author: Lauretta Bender

Range: Age 4 and over

Administration: Individual; 10 minutes; must be administered by a trained examiner

Description: The original Bender-Gestalt is the Visual Motor Gestalt Test by Bender. It consists of 9 designs to be copied by the subject on paper. Various other tests represent adaptations and modifications in administrative procedure, scoring, or expanded interpretive procedures, rather than changes in test material. The most extensive work in adapting the test has been done by Koppitz. The test has been used for a variety of clinical purposes: predictor of school achievement, indicator of emotional disturbance or brain injury, and a measure of visual perception. Bender and others hypothesized that children with normal intelligence, without personality deviations or organic brain disorders, tend to copy designs in a predictable manner. This instrument is used primarily by clinicians and research workers as a diagnostic or experimental technique. Further information is included in Buros Third, Fourth, Fifth, Six, and Seventh Mental Measurements Yearbooks.

Publisher: The Pyschological Corp.
304 East 45th Street
New York, New York 10017

Title: Benton Revised Visual Retention Test

Date: 1974

Author: Arthur L. Benton

Range: Age 8-adult

Administration: Individual; 25 minutes; trained examiner required

Description: The RVRT is a clinical and research instrument designed to measure visual memory, visual perception, and visual-motor coordination. It is used to supplement examinations of persons suspected of abnormality or impairment.

Publisher: The Pyschological Corp.
304 East 45th Street
New York, New York 10017

Title: Boehm Test of Basic Concepts

Date: 1970

Author: Ann E. Boehm

Range: Grades K-2

Adminstration: Group; 30 minutes in 2 sessions; 2 booklets (not equivalent)

Description: The BTBC was designed to assess beginning school children's knowledge of frequently used basic concepts. It is comprised of two booklets, each containing 25 items. The booklets are not intended to be equivalent forms, but are both to be administered. Suggestions are made in the manual on how to use the results to plan remedial instruction. It appears that the test is not of great value for first graders from middle or upper socioeconomic levels and for most second graders (McCandless, 7th MMY). All reviewers in the 7th MMY felt that the BTBC was basically well developed.

Development: The standardization population consisted of 9,700 children enrolled in kindergarten, grade one and grade two in sixteen cities across the U.S. Midyear norms were obtained from children in five cities. The subjects were described by city, grade and socioeconomic level.

Reliability: Split-half reliability coefficients ranging between .68 and .90 for the total score were reported.

Validity: No evidence except "face" validity was reported.

Publisher: Pyschological Corp.
304 East 4th Street
New York, New York 10017

Title: Botel Reading Inventory

Date: 1970

Author: Morton Botel

Range: Pupils in grades 1-12 whose reading levels are fourth grade or lower

Administration: Individual; 40 minutes; 2 forms

Description: The BRI is an informal reading inventory consisting of four subtests designed to aid the teacher in estimating the instructional, independent and frustration reading levels of children and to evaluate knowledge of selected phonics and related skills. The subtests cover word recognition, word opposites, phonics mastery and spelling. The BRI could be useful as an informal assessment; however, the reading teacher could gather as much information from her own informal reading inventory.

Development: The BRI is not a standardized test

Reliability: Reliability was reported for Word Recognition and Word Opposites subtests only.

Validity: Validity was reported for Word Recognition and Word Opposites subtests only.

Publisher: Follett Publishing Company
2020 West Washington Boulevard
Chicago, Illinois 60607

Title: Cain-Levine Social Competency Scale

Date: 1963

Authors: Leo F. Cain, Samuel Levine and Freeman F. Elzey

Range: Ages 5-13

Administration: Individual rating; 10 minutes

Description: The Cain-Levine is a 44-item behavioral rating scale designed to estimate the social competence of trainable mentally retarded children. The respondent should be one who is well acquainted with the child such as his mother, teacher or ward personnel. The items are in four subscales: Self-help, Initiative, Social Skills and Communication. Subscales would be useful in planning and evaluating individual programs.

Development: The standardization population consisted of 700 trainable mentally retarded children in California. The intelligence quotients ranged from 25 through 59, and mental ages ranged from two through seven years. Data on parents' education, income and occupation are provided.

Reliability: Test-retest reliability coefficients for 35 subjects over a period of three weeks reportedly ranged from .88 to .97, with an overall coefficient of .98.

Validity: No data on validity were reported in the manual. A discussion of item selection procedures was included.

Publisher: Consulting Psychologists Press
577 College Avenue
Palo Alto, California 94306

Title: California Short-Form Test of Mental Maturity

Date: 1963

Authors: Elizabeth T. Sullivan, Willis W. Clark and Ernest Tiegs

Range: Grades K-adult; 8 levels

Administration: Group; 35-45 minutes

Description: The CSFTMM is an adaptation of the California Test of Mental Maturity. Both were designed to provide "information about functional capacities that are basic to learning, problem-solving, and responding to new situations." The California tests were patterned after the Stanford-Binet. The complete battery is divided into Language and Non-Language sections. Each section yields an IQ and mental age. Reviews in the Mental Measurements Yearbooks are mixed. However, there seems to be general agreement that it is questionable to use Levels 1 and 2 to evaluate younger children.

Development: Approximately 39,000 subjects from 49 states in the U.S. were tested for the 1963 revision. Consideration was given to geographic region, school population, school grade and community size.

Reliability: Reliability coefficients for the total score were reported as ranging from .91 to .95, with the exception of Level 0 which was .78. Coefficients were also reported for sections and variables.

Validity: Comparisons were made between the CSFTMM and other standardzied tests of mental ability. The correlation coefficients were reported in the technical report.

Publisher:
California Test Bureau
McGraw-Hill Book Company
Del Monte Research Park
Monterey, California 93940

Title: Classroom Reading Inventory

Date: 1969

Author: Nicholas J. Silvaroli

Range: Grades 2-8

Administration: Individual; spelling survey may be group administered; 12-15 minutes; 2 forms

Description: The CRI was designed as a diagnostic tool for the elementary classroom teacher who has no prior experience with group or individual diagnostic tests. The purpose of the CRI is to help the teacher determine the hearing-capacity, independent, instructional, and frustration levels. The CRI consists of graded wordlists (Part I), the graded oral paragraphs (Part II) and spelling survey (Part III). The booklet only is purchased, and the publisher grants permission to reproduce the necessary portions from the booklet. Even though the manual states otherwise, some degree of sophistication would be necessary to interpret the results of this inventory.

Development: No information was given on the development of the inventory with the exception of references used for readability measures.

Reliability: No information was reported in the manual.

Valdity: No information was reported in the manual.

Publisher: William C. Brown Company Publishers
135 South Locust Street
Dubuque, Iowa 52001

Title: Cognitive Abilities Test (Primary Battery)

Date: 1968

Authors: Robert L. Thorndike, Elizabeth Hagen and Irving Lorge

Range: Grades K-3

Administration: Group; 1 hour in two sessions; 2 forms

Description: The CAT is part of an integrated series of group intelligence tests. It can be followed by the multi-level edition of the CAT. Normative links between the two tests make is possible to obtain comparable test scores across 14 grades. The CAT uses pictures and oral directions which eliminate the influence of reading skills. According to Cox (7th MMY) there are few group tests designed to measure cognitive ability in the early grades. He further stated that the practical features of the CAT are impressive but further reliability and validity data are needed.

Development: The normative samples for the CAT were drawn from a cross section of schools which were used in the *Lorge-Thorndike* (180,000 pupils in 40 states).

Reliability: Reliability coefficients were reported as being near .90 at each grade level (samples of 300 at each level).

Validity: No validity data were given, with the exception of a factor analysis of the CAT and grade level 3 of the *Lorge-Thorndike*. This analysis showed that 83 percent of the variance was accounted for by a "general reasoning" factor.

Publisher: Houghton Mifflin
53 West 43rd Street
New York, New York 10036

Title: Cooperative Preschool Inventory

Date: 1970

Author: Bettye M. Caldwell

Range: Ages 3-6½

Administration: Individual; 15 minutes

Description: The CPI was designed to be a brief assessment and screening procedure for young children. The original purpose was to measure a child's concepts and basic information prior to his enrollment at Head Start. The 1970 version is reduced considerably in length from the 1968 version.

Development: The standardization of the revised edition took place in eleven Head Start centers. More subjects were tested between the ages of 4 years and 5 years, 5 months than at either extreme.

Reliability: Reliability coefficients for the age groups reportedly ranged from .86 to .92.

Validity: Validity coefficients between the CPI and the Stanford-Binet were reported ranging from .39 to .65.

Publisher: Educational Testing Service
Cooperative Tests and Services
Box 999
Princeton, New Jersey 08540

Title: Denver Developmental Screening Test

Date: 1970

Authors: William Frankenburg and Josiah Dodds

Range: 2 weeks to 6 years

Administration: Individual; 25 minutes

Description: The DDST was designed to be a simple instrument for the detection of children with serious developmental delays. It may be used by persons with no training in psychological testing. The DDST is composed of 105 items which measure development in four areas: personal-social, fine-motor adaptive, language, and gross motor. According to Werner (7th MMY) the DDST is most useful in the 30-month to 4-year range. The language subtest is possibly unfair to minority group children. The authors of the DDST caution users that the test was not designed as a diagnostic tool, to yield a developmental or mental age, or an intelligence quotient.

Development: A total of 240 items was selected form a survey of more than 12 infant and preschool tests. One hundred five items which met established criteria comprise the final instrument. The 105 items were chosen on the basis of a pretest by four medical students. The standardization group contined 1,000 normal children between the ages of 2 weeks and 6 years, 4 months in Denver, Colorado.

Reliability: Test-retest reliabilities and interexaminer reliabilities reported were based on small samples of children (20 and 12) representing a wide age range (2 months to 5½ years). A tester-observer reliability study of 76 children (4 to 77 months) and a test-retest study (7 days apart) of 186 children (1.5 to 76 months) have been more recently completed. For 13 age groups, coefficients ranged between .66 and .93. Reliabilities were poor at early ages but acceptable in the third and fourth years.

Validity: In a more extensive validity study than the study in the 1968 manual, 236 children were evaluated with the DDST and four criterion tests of intelligence and development. Correlations ranged from .74 to .97.

Publisher:

Ladoca Project and Publishing Foundation, Inc.
East 51st Avenue and Lincoln Street
Denver, Colorado 80216

Title: Developmental Test of Visual-Motor Integration

Date: 1967

Authors: Keith E. Beery and Norman A. Buketenica

Range: Ages 2-15 (Long form, 2-15; Short form, 2-8)

Administration: Group; 10 minutes

Description: The VMI (long form) consists of 24 geometric designs of increasing difficulty which are copies with pencil and paper. The VMI was designed primarily for preschool and early primary levels. Remedial strategies are suggested in the manual. Chissom (7th MMY) stated that the VMI should be substantially more useful than other similar tests when more information is provided regarding reliability and validity.

Development: A group of 1,000 children from Illinois was used to standardize the test. Approximately one half were from suburban schools.

Reliability: A reliability coefficient of .93 was reported for 594 suburban pupils. Test-retest reliability over a two-week period for 171 rural pupils was .83 for boys and .87 for girls.

Validity: A correlation of .89 between the VMI and chronological age was reported. One study using 342 pupils indicated a correlation of .50 between the VMI and first grade reading achievement. Another study using 60 pupils at three grade levels showed that the correlation with mental age decreased from .59 to .38 from the first to the seventh grade.

Publisher: Follett Publishing Company
1010 West Washington Boulevard
Chicago, Illinois 60607

Title:	Devereux Adolescent Behavior Rating Scale
Date:	1967
Authors:	George Spivack, Jules Potts and Peter E. Haimes
Range:	Ages 13-18
Administration:	Individual rating; 15 minutes
Description:	The Devereux was designed for use with diagnosed groups of disturbed children and as a help in identifying disturbed children. It attempts to describe and communicate the "overt behavior symptoms which help define the total clinical picture of disturbed adolescents." It does not appear useful for making fine discriminations among normal children.
Development:	The standardization samples consisted of 550 institutionalized adolescents. Ratings were also obtained for comparison on 400 normal children.
Reliability:	According to Jesness (7th MMY), the relability appears adequate.
Validity:	Empirical data were not available on validity.
Publisher:	Devereux Foundation Press 208 Old Lancaster Road Devon, Pennsylvania 19333

Title: Devereux Child Behavior Rating Scale

Date: 1966

Authors: George Spivack and Jules Spotts

Range: Ages 8-12

Administration: Individual rating; 10 minutes

Description: This Devereux scale was developed to assess the behaviors of emotionally disturbed and mentally retarded children. It provides a profile of problem behaviors which may have led parents or other adults to believe that the child is having a problem that requires professional intervention.

Development: Normative data were gathered on 250 disturbed children in residential treatment centers, 100 retarded children in residential treatment centers, and 350 normal children in public schools.

Reliability: The median coefficient of interscorer reliability was reported as .83. A one-week median test-retest coefficient of .83 was reported. The median factor score reliability coefficient was .91.

Validity: Empirical data were not available on validity.

Publisher: Devereux Foundation Press
208 Old Lancaster Road
Devon, Pennsylvania 19333

Title: Devereux Elementary School Behavior Rating Scale

Date: 1967

Authors: George Spivack and Marshall Swift

Range: Grades K-6

Administration: Individual rating; 10 minutes

Description: The *Devereux Elementary Scale* was developed for use by elementary classroom teachers. It is oriented toward behaviors that interfere with academic achievement so that appropriate remedial and/or preventive action can be taken early in school. According to Littell (7th MMY), a major strength of this scale is the care given to the item selection and groupings. The authors caution that the scale "is not intended to provide a measure of 'personality' or character 'traits' " but it will provide a profile of overt behaviors. Littell states that if users confine themselves to this conservative interpretation, the scale will be a convenient tool.

Development: Normative data were obtained from 13 elementary schools, 32 kindergarten through grade six teachers, with ratings of 800 children. The children were described in terms of age, IQ, parental education and race.

Reliability: Test-retest reliability coefficients (one week interval) ranged from .85 to .91, with a medium coefficient of .87.

Validity: No validity data were available.

Publisher: Devereux Foundation Press
208 Old Lancaster Road
Devon, Pennsylvania 19333

Title: Diagnostic Chart for Fundamental Processes in Arithmetic

Date: 1953

Authors: G.T. Buswell and Lenore John

Range: Grades 2-8

Administration: Individual; 20 minutes

Description: The Diagnostic Chart was designed to help the classroom teacher analyze a pupil's performance in one of the fundamental processes (addition, subtraction, multiplication, or division) to determine what is causing him difficulties. The pupil is given a graded series of examples in each of the four operations. A checklist of habits for each operation is provided so that the teacher can record observations. Remedial procedures are suggested in the manual. This instrument is not standardized, nor is it intended for classification purposes. The results are meant to help the teacher assess the child's performance in the basic computation skills.

Development: The items were given to 500 children in Chicago to arrange in order of difficulty. The experimental form was used in 70 classrooms in ten schools in Chicago and its suburbs.

Reliability: No information was reported.

Validity: No information was reported.

Publisher: The Bobbs-Merrill Company, Inc.
4300 West 62nd Street
Indianapolis, Indiana 46206

Title: Diagnostic Reading Scales

Date: 1963

Author: George D. Spache

Range: Grades 1-8; retarded readers in grades 9-12

Administration: Individual; 40 minutes

Description: The DRS is comprised of word recognition tests, paragraphs of graduated difficulty and six phonics tests. The reading passages are designed to yield instructional, independent and potential reading levels. In a review by Barr (7th MMY) she states that the scales are diagnostically useful and appear most valid for first to fourth grades. Comparisons among instructional, independent and potential reading levels cannot be justified on the basis of standardization information which is in the manual. She also advises against using the grade norms for the phonics tests. According to Bryant (6th MMY), this test provides a meaningful approach toward the diagnosis of reading skills and difficulties. Sophisticated interpretation is required, however.

Development: According to Barr, "The instrument is beautifully conceived but standardization leaves much to be desired." No information was presented which describes subjects used as a sample for the development of the scales.

Reliability: Reliabilities of .84 to .88 for the reading passages and .87 to .96 for the word recognition lists were reported.

Validity: Validity coefficients are reported for grades 2 through 6 and three variables of the scales. These ranged from .64 to .92.

Publisher: California Test Bureau/McGraw-Hill
Del Monte Research Park
Monterey, California 93940

Title: Diagnostic Tests and Self-Helps in Arithmetic

Date: 1955

Author: Leo J. Brueckner

Range: Grades 3-12

Administration: Group; no time limits

Description: This series includes four screening tests and 23 diagnostic tests. The screening tests examine whole numbers, fractions, decimals and general arithmetic skills. Each of the diagnostic tests assesses a different skill, and each is accompanied by a self-help exercise. There are no norms for these tests, nor is there information on reliability and validity. The tests are intended to be instructionally helpful in determining specific strengths and weaknesses of individual children.

Publisher: California Test Bureau
McGraw-Hill
Manchester Road
Manchester, Missouri 63011

Title:	Doren Diagnostic Reading Test
Date:	1973
Author:	Margaret Doren
Range:	Grades 1-4 or higher for diagnosis of remedial students
Administration:	Group; 1 to 3 hours depending on size and level of group
Description:	The DDRT is designed to provide an objective measure of basic word recognition skills. Its purpose is to determine a starting point for remedial instruction. The reading techniques covered are those usually taught in the primary grades. The skills covered are letter recognition, beginning sounds, whole word recognition, words within words, speech consonants, ending sounds, blending, rhyming, vowels, discriminate guessing, spelling and sight words. The skills are plotted on a skill profile sheet for each pupil. Remedial activities are included in the manual. According to Feldmann (1974), this test has flaws in construction and administration procedures which seem to limit its usefulness for the teacher.
Development:	No information was given in the manual on item selection except that items were "selected from a larger number of tested items." The test was standardized on a population of 165 first through fourth graders in four unidentified suburban school districts in the Midwest.
Reliability:	No information was reported.
Validity:	A correlation of .90 between reading achievement and the DDRT was reported for eleven classes in grades one to four.
Publisher:	American Guidance Service, Inc. Publisher's Building Circle Pines, Minnesota 55014

Title:	Durrell Analysis of Reading Difficulty
Date:	1955
Author:	Donald D. Durrell
Range:	Grades 1-6
Administration:	Individual; untimed, approximately 30-60 minutes
Description:	The Durrell consists of a series of tests and situations during which the examiner may observe in detail various aspects of a child's reading. The primary purpose is to observe faulty habits and weaknesses in reading in order to plan a remedial program. The following areas are assessed: silent and oral reading, listening comprehension, word analysis, phonetics, pronunciation, writing and spelling. According to Robinson (4th MMY), this test provides the experienced reading teacher an excellent opportunity to observe difficulties in word recognition and oral reading. She further stated that the norms must be used with full realization of their limitations.
Development:	Norms are given for some of the subtests. The only information provided is that if norm tables are presented, the norms are based on a minimum of 1,000 children for each test.
Reliability:	No information was reported.
Validity:	No information was reported.
Publisher:	Harcourt Brace Jovanovich 757 Third Avenue New York, New YOrk 10017

Title: Durrell Listening-Reading Series

Date: 1970

Authors: Donald D. Durrell, Mary T. Hyes and Mary B. Brassard

Range: Grades 1-9; higher for remedial students; 3 levels (primary, intermediate and advanced)

Administration: Group; 1½-2 hours; 2 forms

Description: This series consists of a reading test which measures reading achievement and a listening test which measures understanding of the spoken word. It is designed to measure the degree of retardation in reading as compared with listening. This series replaces Durrell-Sullivan Reading Capacity and Achievement Tests (1944). According to Bormuth (7th MMY), in spite of some weakness in design of test items, each of these reading and listening tests is "useful in its own right and compares very favorably with other available tests of reading and listening abilities."

Development: Standardization was completed on a population of 20,000 pupils.

Reliability: According to Spache (7th MMY), reliability data presented were adequate for the direct comparisons recommended by the authors.

Validity: Both Bormuth (7th MMY) and Spache (7th MMY) felt content validity was good. Construct and concurrent validity were as yet unsubstantiated.

Publisher: Harcourt Brace Jovanovich
757 Third Avenue
New York, New York 10017

Title: First Grade Screening Test

Date: 1969

Authors: John E. Pate and Warren W. Webb

Range: Grades K-1

Administration: Group; untimed, approximately 45 minutes for kindergarten and 30 minutes for first grade.

Description: The FGST was developed to identify those children who will not make sufficient progress during the first year of school to be ready for second grade. It was designed to identify children who should be referred for further diagnosis and appropriate remedial planning. Separate test booklets are provided for boys and girls. The difference in the two forms is that the sex of the pictured children is the same as the subjects.

Development: The FGST was standardized on a geographically stratified sample of 5,500 first grade students enrolled in 248 classrooms, and 3,200 kindergarten students enrolled in 160 classrooms. The students attended schools ranging from rural one-room schools to large schools in a megapolis.

Reliability: Test-retest reliability analysis based on two-week and eight-week retesting resulted in correlations of .84 and .82. Inter- and intra-scorer reliability of .98 was reported.

Validity: According to the manual, predictive validity was established by obtaining teacher ratings and achievement test criteria later in the year. A pretermined cutoff score identified 68% of the "not ready" group and misidentified only 13% of the "ready" group. Comparisons of FGST scores with various achievement tests resulted in correlations ranging from .60 to .79.

Publisher:
American Guidance Service, Inc.
Publisher's Building
Circle Pines, Minnesota 55014

Title: Gates-McKillop Reading Diagnostic Test

Date: 1962

Authors: Arthur I. Gates and Anne S. McKillop

Range: Grades 2-6

Administration: Individual; 30-60 minutes; 2 forms

Description: The Gates-McKillop is a battery of 17 tests which involve oral responses to allow the examiner to analyze specific areas of reading difficulty. Not all children need to be administered all 17 subtests. Comprehension is not measured. According to Bryant (6th MMY) this test requires a substantial amount of sophisticated clinical judgment to interpret a student's performance.

Development: Although norms were provided in the manual, no information was given about the standardization population on which the norms were based.

Reliability: No information was reported.

Validity: No information was reported.

Publisher: Teachers College Press
1234 Amsterdam Avenue
New York, New York 10027

Title: Gates-MacGinitie Reading Tests

Date: 1965

Authors: Arthur I. Gates and Walter H. MacGinitie

Range: Grades 1-12

Administration: Group; 50 minutes in two sessions

Description: This series replaces the *Gates Primary Reading Tests* and *Gates Reading Survey*. The test consists of three parts designed to measure reading achievement: Speed and Accuracy, Vocabulary and Comprehension. The content and nature of the items are essentially the same as the earlier version. The *Gates-MacGinitie* tests are best used to survey reading achievement and yield little diagnostic information.

Development: Test items were selected on basis of a nationwide sample of 25,000 pupils. The tests were standardized on 40,000 pupils in 38 communities selected on basis of size, location, average educational level and average family income.

Reliability: Alternate-form and split-half reliabilities were reported.

Validity: No information was reported in the manual.

Publisher: Teachers College Press
1234 Amsterdam Avenue
New York, New York 10027

Title: Gates-MacGinitie Reading Tests: Readiness Skills

Date: 1969

Authors: Arthur I. Gates and Walter H. MacGinitie

Range: Grades K-1

Administration: Group; 120 minutes in 4 sessions

Description: The purpose of the Gates-MacGinitie is not clearly stated in the manual (Sykstra, 7th MMY). It is assumed that the test was designed to measure reading readiness rather than general school readiness. Eight subtests are included: Listening Comprehension, Auditory Discrmination, Visual Discrimination, Following Directions, Letter Recognition, Visual-Motor Coordination and Auditory Blending. According to Dykstra, the test may be an excellent one as it is the joint product of two highly respected authorities. However, limitations in the manual present a judgment based on the information presented.

Development: The standardization population consisted of 4,500 children in 35 communities. The subjects were selected on the basis of community size and location, educational level of parents and family income.

Reliability: The technical supplement reported subtest reliabilities but no reliability data for the total test.

Validity: Correlations between the readiness score and the Gates-MacGinitie Reading Test were presented. For first graders the correlations between the total readiness scores and the vocabulary and comprehension scores of the reading test were .60 and .59.

Publisher: Teachers College Press
1234 Amsterdam Avenue
New York, New York 10027

Title: Gilmore Oral Reading Test

Date: 1968

Authors: John V. Gilmore and Eunice C. Gilmore

Range: Grades 1-8

Administration: Individual; 15 minutes; 2 forms

Description: The Gilmore was designed to measure three aspects of oral reading: accuracy, comprehension and rate. The Accuracy Subtest score measures fluency and accuracy of word recognition. Smith (7th MMY) questioned the basic assumption of the test - that comprehension should be measured with an oral test. It would seem more appropriate to assess comprehension through a silent reading task. Comprehension questions seem to measure short-term recall of details, rather than requiring inferences or judgments to be made. According to A.J. Harris (7th MMY), the Gilmore is among the best standardized tests of accuracy in oral reading of meaningful material. He questions the usefulness of the Comprehension and Rate scores.

Development: The standardization population included 1,600 pupils from five states in grades 1-8.

Reliability: Reliability coefficients for the Accuracy scores ranged from .84 to .94. For Comprehension and Rate scores, the coefficients ranged from .53 to .70.

Validity: Statistical evidence of validity was obtained from a comparison of the scores of 24 fifth grade pupils in this test, the Gray Oral Reading Test and the oral reading test from the Durrell Analysis of Reading Difficulty. Correlations ranged from .39 to .80.

Publisher: Harcourt Brace Jovanovich
757 Third Avenue
New York, New York 10017

Title: Goldman-Fristoe-Woodcock Auditory Skills Test Battery

Date: 1975

Authors: Ronald Goldman and Macalyne Fristoe

Range: Age 3 and over

Administration: Individual; 10-15 minutes for each of 12 tests

Description: The G-F-W Battery is designed to provide a wide range of diagnostic instruments for use in identifying subjects who are deficient in auditory skills and to provide information describing these deficiencies. The G-F-W Battery consists of 12 tests in five easel kits. Major areas covered are Selective Attention, Auditory Discrimination, Auditory Memory and Sound-Symbol.

Development: The standardization population consisted of 7,200 normal and clinic subjects in four geographic areas. Norms for the subtest, Reading of Symbols, were based on 5,250 different subjects.

Reliability and Validity: The technical manual for the G-F-W Battery is to be available in November, 1975. At the time this summary was prepared, no information on reliability and validity was available. The test should be used and interpreted cautiously until empirical evidence is available.

Publisher: American Guidance Service
Publisher Building
Circle Pines, Minnesota 55014

Title: Goldman-Fristoe-Woodcock Test of Auditory Discrimination

Date: 1970

Authors: Ronald Goldman, Macalyne Fristoe and Richard W. Woodcock

Range: Ages 4 and over

Administration: Individual; 15 minutes

Description: The G-F-W was designed for the purpose of identifying and assessing the subject's ability to distinguish among speech sounds. The measurement takes place during both quiet and noisy conditions by means of prerecorded tapes. According to Sheeley (7th MMY), the G-F-W is more suitable than other auditory tests for use with very young subjects or those with verbal communication defects. This is because it incorporates a pointing response, more than two possible choices for each item, and pretraining to teach vocabulary.

Development: The standarization sample consisted of 745 apparently normal hearing subjects, ages 3 to 84. The subjects were from Minnesota, New Jersey and Tennessee.

Reliability: Test-retest reliabilities of .87 (Quiet Subtest) and .81 (Noise Subtest) were reported for a very small sample of 17 preschool, speech-handicapped children. Split-half reliability coefficients were reported as ranging from .51 to .88 (Quiet Subtest) and .63 to .68 (Noise Subtest).

Validity: The G-F-W appears to discriminate among the performances of various "clinical" groups and the performance of "normals." According to Proger (7th MMY), " . . . the G-F-W appears to be one of the few special education instruments with reasonably adequate research and development. Concerns about validity, et. al., plague all the tests in the field. The authors have made more of an effort than most to deal with them."

Publisher:
American Guidance Service
Publisher's Building
Circle Pines, Minnesota 55014

Title:	Gray Oral Reading Test
Date:	1967
Authors:	William S. Gray; edited by Helen M. Robinson
Range:	Grades 1-12 and adults
Administration:	Individual; time not reported; 4 forms
Description:	The Gray was designed to measure oral reading growth and to aid in the diagnosis of oral reading difficulties. It is suggested that it be used in conjunction with a measure of silent reading. Thirteen passages of increasing levels of difficulty are provided. According to A.J. Harris (6th MMY) and Lohnes (6th MMY) this test is a welcome addition to available oral reading tests.
Development:	The standarization population was based on results obtained by administering all four forms of the test to 500 pupils in grades 1-12. Subjects were from schools in Florida and Chicago. Attempts were made to have "average readers" for each grade in the norming population. The average level of subjects tended to be toward the upper limit of average.
Reliability:	Alternate-form reliability coefficients ranged from .97 to .98.
Validity:	No information was available.
Publisher:	Bobbs-Merrill Company, Inc. 4300 West 62nd Street Indianapolis, Indiana 46268

Title: Hiskey-Nebraska Test of Learning Aptitude

Date: 1966

Author: Marshall S. Hiskey

Range: Ages 3-16

Administration: Individual; 45 minutes; trained examiner

Description: The HNTLA was originally developed as a test of learning aptitude of the deaf. It was revised in 1955 to include norms and instructions for hearing children. However, it probably is more frequently used with deaf subjects as other suitable measures are available for hearing children. According to Newland (7th MMY), this test provides a better measure of components necessary for school success by deaf children, especially younger ones, than do other tests. According to the author of the test, the items involve tasks similar to those which a deaf child must perform in school.

Development: The standardization population consisted of 1,100 deaf children and 1,100 hearing children, ages 2 years, six months to 17 years, 5 months, in ten states. No breakdowns by race or sex were provided. Level of parental occupation corresponds closely to U.S. census data.

Reliability: Split-half reliability coefficients were reported as .95 for deaf and .93 for hearing subjects in groups between the age range of 3 to 10; .92 and .90 were reported for groups in the range of 11 to 17 years.

Validity: For the deaf, subtest intercorrelations were reported as ranging from .33 to .74 for the ages 3 to 10, and from .31 to .43 for the ages 11 to 17. A correlation of .86 with the Stanford-Binet for subjects ranging in age from 3 to 10, and .78 for subjects ranging in age from 11 to 17. A correlation of .82 with the WISC was reported.

Publisher: Marshall S. Hiskey
5640 Baldwin
Lincoln, Nebraska 68508

Title: Illinois Test of Psycholinguistic Abilities

Date: 1968

Authors: Samuel A. Kirk, James J. McCarthy and Winifred D. Kirk

Range: Ages 2-10

Administration: Individual; 45-60 minutes; trained examiner

Description: The ITPA was constructed to assess the psychological functions through which an individual communicates with others and through which he receives communication from others or from the enviornment. The ITPA was designed to be primarily diagnostic in nature. Through 12 subtests, it evaluates abilities in the visual-motor and auditory-vocal channels of communication. Two levels of organization (representational and automatic) are recognized. Within these levels three psycholoiguistic processes (receptive, organizational and expressive) are evaluated. Various researchers have cautioned that there are certain limitations of the ITPA which should be carefully considered:

1. At the extremes of the age range for which it was developed, the ITPA results are questionable.
2. Use of the ITPA with lower-class children (especially minority groups), adolescents, and adult retardates cannot be justified.
3. Some examiners attempt to use the psycholinguistic quotient (PLQ) as a measure of intelligence. This is not a limitation of the test, but of an unqualified examiner.
4. Factorial composition, reliability and validity of the Revised Edition still need to be explored.

Development: The standardization population for the Revised Edition consisted of 962 children, ages 2 through 10. The group was from homes slightly above the national average in income and education. Only children who had average intelligence, average school achievement, average adjustment, and who had no gross sensory-motor deficits were included. The sample was from five Midwestern towns of moderate size. Metropolitan and rural areas were not included.

Reliability: According to Chase (7th MMY) the subtests are reasonably reliable at each age level with the exception of Visual Closure and Auditory Closure. Hammill and Bartel (1975) stated that all subtests have reliabilities high enough to support their clinical use except the Visual Closure Subtest.

Validity: The factorial composition of the Revised Edition is still being explored. One study identified only three interpretable factors in the 1961 version (Uhl and Murss, 1970). These factors were vocabulary, immediate memory span and auditory precessing. Newcomer and her colleagues (1975) supported the construct validity of the ITPA. According to Carroll (7th MMY), "the authors (of the ITPA) have an obligation to report further research supporting the usefulness of the ITPA for the major purpose for which it was designed - diagnosis and remediation."

Publisher: University of Illinois Press
Urbana, Illinois 61801

Title: KeyMath Diagnostic Arithmetic Test

Date: 1971

Authors: Austin J. Connolly, William Nachtman and E. Milo Pritchett

Range: Preschool-grade 6; no upper limit for individual remedial use

Administration: Individual; 30 minutes

Description: The *KeyMath* is designed to provide a diagnostic assessment of mathematic skills beyond the four fundamental computation processes. The 14 subtests are organized into three major areas: Content, Operations, and Applications. Each subtest contains items arranged in order of increasing difficulty. Every pupil is not administered all items.

Development: Items were selected on the basis of administration to 951 subjects grades K-8. Normative data were gathered on 1,200 subjects in grades K-7. The subjects were randomly selected from 42 schools in eight states.

Reliability: Total test score reliabilities were reported for each grade from K-7. These correlations ranged from .94 to .97.

Validity: The content and sequence were based on analyses of ten major mathematics programs. No data were reported on the validity of the final form of *KeyMath*.

Publisher: American Guidance Service, Inc.
Publisher's Building
Circle Pines, Minnesota 55014

Title: Kindergarten Auditory Screening Test

Date: 1971

Author: Jack Katz

Range: Grades K-1

Administration: Groups of 6-10; 20 minutes

Description: The purpose of the KAST is to identify kindergarten and first-grade children who have "auditory perception difficulties." KAST is a screening device on record and contains three subtests for assessing auditory skills. The subtests involve sound blending, figure ground and auditory discrimination skills.

Development: The standardized population consisted of 600 kindergarten and first-grade pupils in five states (Massachusetts, Oregon, Texas, Illinois and Missouri).

Reliability: No information was reported in the manual.

Validity: No information was reported in the manual.

Publisher: Follett Publishing Company
1010 West Washington Boulevard
Chicago, Illinois 60607

Title: Lee-Clark Reading Readiness Test

Date: 1962

Authors: J. Murray Lee and Willis W. Clark

Range: Grades K-1

Administration: Group; 20 minutes

Description: The LCRRT is one of the first tests designed to help determine which pupils are ready for reading instruction. It consists of three subtests: Letter Symbols, Concepts and Word Symbols. One advantage of this test is that considerably less testing time is required than for most readiness tests. Berg (7th MMY) gave the LCRRT a generally good overall rating with the qualification that other measures of readiness should be used to evaluate pupil readiness with the LCRRT.

Development: Norms were based on 5,000 entering first graders with a median CA of 6 years and a median IQ of 100. Norms are also available for the end of kindergarten.

Reliability: Split-half reliabilities were reported ranging from .87 to .96.

Validity: Predictive validity coefficients were reportedly in the .40's and .50's.

Publisher:
California Test Bureau
Division of McGraw-Hill
Manchester Road
Manchester, Missouri 63011

Title: Lincoln-Oseretsky Motor Development Scale

Date: 1965

Author: Adapted by William Sloan

Range: Ages 6-14

Administration: Individual; untimed

Description: The *Lincoln-Oseretsky* is an adaptation of the *Oseretsky Test*, developed in Russia. Other versions of the *Oseretsky* are also in use. It is designed to assess gross-motor skills. The *Lincoln-Oseretsky* consists of 36 items arranged in order of difficulty. The skills involved include speed, dexterity, coordination, rhythm, balancing, and jumping. A motor age is calculated, based on performance.

Development: The normative data were from a sample of 750 boys and girls, ages 6-14, from small towns in Central Illinois. The manual cautions that the norms should be considered tentative as the sample size was limited, and the subjects were chosen mainly for their availability.

Reliability: Split-half reliability coefficients at each age level ranged between .72 and .94 for males and .82 to .93 for females. The one exception was for 14-year-old females where the coefficient was .59. The median reliability coefficient was .86. These coefficients were computed for the sample of 750 pupils.

Validity: Low correlations were reported between the *Lincoln-Oseretsky* and other tests of motor ability: the *Brace Scale of Motor Ability* (.32), the *Cowan Pratt* test (.37), and the *Metheny-Johnson* test.

Publisher: Western Psychological Services
12031 Wilshire Boulevard
Los Angeles, California 90025

Title: Lorge-Thorndike Intelligence Tests, Multi-Level Edition

Date: 1966

Authors: Irving Lorge, Robert L. Thorndike and Elizabeth P. Hagen

Range: Grades 3-13; 8 levels

Administration: Group; Verbal Battery, 35 minutes; Nonverbal Battery, 30 minutes; 2 forms

Description: The Multi-Level Edition is an outgrowth of an earlier Separate Level Edition of the L-T Tests. The purpose of the L-T Tests is to provide an estimate of mental ability independent of reading ability. The Verbal Battery is composed of five subtests using only verbal items. The Nonverbal Battery uses items which are either pictorial or numerical. According to Tittle (7th MMY), the Multi-Level Edition is a refinement and improvement over the earlier edition. The L-T Tests meet generally accepted standards for test construction and standardization procedures.

Development: The L-T Tests were restandardized in 1963. This standardization was carried out with the *Iowa Tests of Basic Skills* (grades 3-8) and the *Tests of Academic Progress* (grades 9-12). According to Tittle (7th MMY), the standardization procedures generally appeared to have been carefully carried out.

Reliability: Alternate-form reliability coefficients were reported as .83 to .91 (Verbal Battery) and .80 to .88 (Nonverbal Battery) for the various grade levels.

Validity: Correlations between the L-T Tests and tests of achievement were computed. These coefficients ranged in the .60's and .70's, with some in the .80's. Moderate correlations were reported with other intelligence tests. One series of studies indicated that the Verbal Battery is more predictive of school achievement than the Nonverbal Battery.

Publisher: Houghton Mifflin
1900 South Batavia Avenue
Geneva, Illinois 60134

Title: Lorge-Thorndike Intelligence Test, Primary Battery

Date: 1957

Authors: Irving Lorge and Robert L. Thorndike

Range: Grades K-1; 2 levels

Administration: Group; 3 sessions, 10 minutes each

Description: The L-T, Primary Battery is part of the earlier L-T, Separate Level Edition. The Primary Battery uses pictorial type items; thus, the poor reader is not penalized. The tests are untimed, with the examiner determining the speed of administration. The manual suggests that the results may be used, in combination with other tests, to form class groups, to group within a class, or to set standards of expectancy.

Development: The norms were based on stratified community samples. Over 136,000 children in 44 communities in 22 states were tested. Four norms were developed: IQ equivalent, grade percentile, grade equivalent, and age equivalent.

Reliability: Standard errors of measurement were reported in the manual, rather than reliability coefficients.

Validity: The Primary Battery reportedly correlated .56, .63 and .69 with three unspecified "well-known group test of intelligence."

Publisher: Houghton Mifflin
1900 South Batavia Avenue
Geneva, Illinois 60134

Title: Marianne Frostig Developmental Test of Visual Perception

Date: 1966

Authors: Marianne Frostig in collaboration with Welty Lefever and John R.B. Whittlesey

Range: Ages 3-8

Administration: Group; 40-60 minutes

Description: The DTVP was designed to measure five perceptual skills: eye-motor coordination, figure-ground discrimination, form constancy, position in space, and spatial relations. These were selected by the authors because of these skills' presumed relationship to academic performance. The authors strongly advise only trained examiners administer the test. Mann (7th MMY) cautions users against assuming that low scores on the DTVP indicate a need for perceptual training.

Development: The DTVP was standardized on 2,100 subjects from southern California schools. Low socioeconomic and minority groups are poorly represented.

Reliability: Test-retest reliability of the perceptual quotient was reported as .80 for 72 first and second graders tested by trained examiners two weeks apart. Subtest reliabilities ranged from .42 to .80.

Validity: Correlations between DTVP and first grade reading achievement were reported as .40 to .50. Various validation studies do not support Frostig's five types of perceptual abilities; instead, they demonstrate one "perceptual" factor.

Publisher: Consulting Psychologists Press
577 College Avenue
Palo Alto, California

Title: McCullough Word-Analysis Tests

Date: 1963

Author: Constance M. McCullough

Range: Grades 4-6

Administration: Group; 70 minutes in 7 sessions

Description: The MWAT is a battery of seven subtests designed to measure certain phonic and structural analysis skills. One word recognition technique typically emphasized in the grades 4-6, using context clues, is not measured. The MWAT received a generally favorable review by L.A. Harris (7th MMY). He felt the test would be a useful diagnostic device to be used with other information, such as analysis of comprehension skills. The MWAT was also favorably reviewed by Bliesmer (6th MMY).

Development: The standardization population consisted of pupils in grades 4-6 who had been taught using the Ginn Basic Readers. Approximately 1,800 pupils in 23 school systems in 21 states participated. An attempt was made to control sex, socioeconomic class and intelligence.

Reliability: Reliability coefficients reportedly ranged from .83 to .96.

Validity: No information on validity was available.

Publisher: Personnel Press
191 Spring Street
Lexington, Massachusetts 02173

Title: Meeting Street School Screening Test

Date: 1969

Authors: Peter K. Hainsworth and Marian L. Siqueland

Range: Grades K-1; Ages 5 to 7 years, 5 months

Administration: Individual; 15-20 minutes

Description: The MSSST was designed as a screening battery to survey gross motor, visual-perceptual-motor and language skills. The administration results in 3 subtest scores and a total score. An arbitrary cutoff point is assigned to identify children who may have later learning difficulties.

Development: The MSSST was standardized on 500 children selected to be representative of the U.S. population in age, sex and father's occupation. The majority of the children were from East Providence, Rhode Island.

Reliability: Test-retest reliability (two to four weeks apart) was .85. Inter-rater reliability was .95. The number of children was not given for either figure.

Validity: According to one review by Yule (7th MMY), the validity data are inadequate and difficult to interpret. The total MSSST score correlated .77 with the ITPA and .57 with Frostig's DTVP. No correlation between IQ and the MSSST is reported. According to Yule, it has not been shown that the MSSST measures any areas not measured by Wechsler Scales; and until further studies are completed, the MSSST cannot be accepted as a valid predictor of learning difficulties.

Publisher: Crippled Children and Adults of Rhode Island, Inc.
Meeting Street School
333 Grotto Avenue
Providence, Rhode Island 02906

Title: Metropolitan Readiness Tests

Date: 1969 (An extensive revision is planned for release in January, 1976)

Authors: Gertrude H. Hildreth, Nellie L. Griffiths and Mary E. McGauvran

Range: Grades K-1

Administration: Group; 65-75 minutes in 3 sessions; 2 forms

Description: The MRT is probably the most widely used test to measure readiness for first-grade instruction. It has received generally favorable reviews by most critics. The authors recommend that little significance should be attached to subtest socres of individual pupils. According to Singer (7th MMY), reliability is sufficiently high to use the total score with individuals.

Development: The standardization population consisted of 15,000 first grade students in 70 school systems. The sample is described regarding sex, age, intelligence, community size, geographic distribution and other community variables. Racial or ethnic group characteristics are not mentioned.

Reliability: An alternate-form reliability coefficient of .91 was reported for the total test. Subtest coefficients were reportedly lower (.50 to .86).

Validity: The manual contains an extensive discussion of the content, construct and predictive validity of the MRT.

Publisher: Harcourt Brace Jovanovich
7555 Caldwell
Chicago, Illinois 60648

Title: Motor-Free Visual Perception Test

Date: 1972

Authors: Ronald P. Colarusso and Donald D. Hammill

Range: Ages 5-8

Administration : Individual; 10 minutes

Description: The MVPT is a test of visual perception which avoids motor involvement in that the child points to a stimulus rather than being required to reproduce it. Five types of visual perception are assessed: spatial relationships, visual discrimination, figure-ground, visual closure and visual memory. The MVPT contains 36 items.

Development: The MVPT was standardized on a sample of 881 normal children ages four through eight from twenty-two states. Subjects from all races, economic levels, and residential areas were included.

Reliability: Test-retest reliability coefficients ranged from .77 to .83 at five age levels with an overall coefficient of .81. Split-half coefficients ranged between .81 and .84 with an overall coefficient of .88. Kuder-Richardson 20 coefficients ranged from .72 to .82 with an overall coefficient of .86.

Validity: The MVPT correlated .73 with the *Frostig*. A median correlation of .31 with two intelligence tests was reported. A median correlation of .38 with tests of school performance was reported.

Publisher:
Academic Therapy Publications
1539 Fourth Street
San Rafael, California 94901

Title: Murphy-Durrell Reading Readiness Analysis

Date: 1965

Authors: Helen A. Murphy and Donald D. Durrell

Range: Grade 1

Administration: Group; 80 minutes in two sessions

Description: The MDRRA is a revision of the Murphy-Durrell Diagnostic Reading Readiness Test (1949). Three subtests are included: phonemes, letter names and learning rate. According to Barr (7th MMY), this test is well constructed and well standardized. Singer (7th MMY) stated that although the MDRRA has some weaknesses, it is still one of the best tests of reading readiness. Teachers can use information from the results to adapt reading instruction to the individual's mode and rate of learning.

Development: The norms are based on 12,200 beginning first graders in 65 schools in 12 states. Normative data are not provided for kindergarten children so local norms would be needed if a school assessed reading readiness in kindergarten.

Reliability: Split-half reliabilities were based on 200 children randomly selected from the standardization sample. These ranged from .88 to .97 for the subtests. Total test reliability was reportedly .98.

Validity: The MDRRA correlated .80 with the Metropolitan Readiness Test. The predictive validity coefficients ranged from .65 to .66 using the Stanford Achievement Test: Reading.

Publisher: Harcourt Brace Jovanovich
7555 Caldwell
Chicago, Illinois 60648

Title: New Developmental Reading Tests (Intermediate)

Date: 1968

Authors: Guy L. Bond, Bruce Balow and Cyril L. Hoyt

Range: Grades 4-6

Administration: Group; 50 minutes; two forms

Description: The NDRT for the intermediate grades consist of five parts: basic reading vocabulary, reading for information, relationships, interpretation, and appreciation. A single comprehension score is obtained from Parts II, III, IV and V. According to Traxler (7th MMY) the intermediate manual shows considerable improvement over the primary manual. Traxler recommended the NDRT for "cautious use" in analyzing reading abilities of individual pupils.

Development: Approximately 15,000 pupils from over 100 schools were included in the standardization sample. Geographical location and community type, size, and socioeconomic level were considered.

Reliability: Alternate-form reliabilities ranging from .77 to .91 were reported from approximately 400 pupils in grades 4 and 6. Internal consistency reliabilities were reported from .80 to .94 for approximately 1,150 pupils.

Validity: The manual contains a discussion of content and construct validity.

Publisher:
Lyons and Carnahan
Rand McNally
407 E. 25th Street
Chicago, Illinois 60616

Title: New Developmental Reading Tests (Primary)

Date: 1965

Authors: Guy L. Bond, Bruce Balow and Cyril L. Hoyt

Range: Grades 1-3; 2 levels

Administration: Group; 55 minutes in 2 or 3 sessions; two forms

Description: The NDRT for the primary grades consists of two levels (lower primary and upper primary). The primary battery contains three parts: word recognition, comprehending significant ideas and comprehending specific instructions. The NDRT is designed to measure general reading growth. According to Davis (7th MMY) and Traxler (7th MMY), the primary tests lack adequate technical data, and there are shortcomings in the manual of instructions.

Development: Approximately 5,000 children were tested from stratified, randomized samples from two large Midwestern communities.

Reliability: Alternate-form reliabilities were reported ranging from .89 to .95. These were based on 150 pupils in five first-grade classrooms.

Validity: No information was reported.

Publisher: Lyons and Carnahan
Rand McNally
407 E. 25th Street
Chicago, Illinois 60616

Title: Northwestern Syntax Screening Test

Date: 1971

Author: Laura Lee

Range: Ages 4-12

Administration: Individual; 20-30 minutes

Description: The NSST is intended as a screening device for use by speech clinicians who need a quick estimate of a child's syntactic development. The test manual states that the test should be used only for screening and only with speakers of the standard English dialect.

Development: The norms are based on a sample of 344 children from middle- to upper-income communities where the standard English dialect is spoken.

Reliability: No information was available.

Validity: No information was available.

Publisher: Northwestern University Press
Evanston, Illinois 60201

Title:	Otis-Lennon Mental Ability Test
Date:	1967
Authors:	Arthur S. Otis and Roger T. Lennon
Range:	Grades K-12; 6 levels
Administration:	Group; 30-60 minutes depending upon level; 2 forms
Description:	The Otis-Lennon is a new edition of the Otis Quick-Scoring Mental Ability Tests. It provides for the assessment of "general mental ability, or scholastic aptitude." The test is primarily a measure of verbal ability. No reading is required on the first three levels. Milholland, Smith, and Grotelueschen, who each reviewed the Otis-Lennon in the Seventh Mental Measurements Yearbook all agree that this test is a better than average, possibly outstanding, test of its kind. The manual cautions interpreting results for children who do not have normal backgrounds and motivation.
Development:	National norms are based upon a sample of 200,000 pupils selected as representative. The sample was controlled for size and type of school, family income and educational level, geographic location and quality of school in terms of educational achievement within its own system.
Reliability:	Alternate-form, test-retest and split-half reliabilities were reported for each grade. According to the three reviewers, substantial evidence which indicates high reliability is provided.
Validity:	Validity is discussed in content, construct and criterion-related categories. Data were presented which supported the use of the Otis-Lennon.
Publisher:	Harcourt Brace Jovanovich 7555 Caldwell Chicago, Illinois 60648

Title: Ohio Tests of Articulation and Perception of Sounds

Date: 1973

Author: Ruth Beckey Irwin

Range: Ages 5-8

Administration: Individual; 15-30 minutes

Description: The OTAPS were designed to evaluate articulation and perception of vowel and consonant sounds. More than one sound can be tested in each word or nonsense syllable. The child's spontaneous and imitative production of sounds, as well as identification and perception of sounds, are assessed in eight subtests.

Development: The standardization population consisted of 200 children, ages 5-8, from Ohio. The subjects were controlled for age, sex, hearing, intelligence and parental occupation level.

Reliability: Intra-rater and inter-rater reliabilities were reported and appeared adequate. Test-retest reliability coefficients for each of the subtests were reported for a small sample of 20 first- and second-grade children. The time period between tests was one week. The coefficients ranged from .24 to .89.

Validity: Validity coefficients were determined for the four subtests on articulation by comparing them with the *Templin-Darley Screening Test of Articulation*. The coefficients reportedly ranged from .87 to .93.

Publisher: Stanwix House, Inc.
3020 Chartiers Avenue
Pittsburgh, Pennsylvania 15204

Title: Peabody Individual Achievement Test

Date: 1970

Authors: Lloyd M. Dunn and Frederick C. Markwardt, Jr.

Range: Grades K-adult

Administration: Individual; 30 minutes

Description: The PIAT was designed for the purpose of providing a wide-range scrcening measure of achievement in the areas of mathematics, reading recognition and comprehension, spelling and general information. It results in a quick, rough estimate which may indicate the need for a more complete diagnosis in one or more areas. French (7th MMY) and Lyman (7th MMY) both agree that the PIAT could be very helpful as a screening instrument which could then be supplemented, if necessary, with a more reliable and thorough test.

Development: The PIAT was standardized on 2,900 subjects from grades K-12. The sample was similiar in characteristics to the total U.S. population in terms of race, sex, age parental occupation and type of community.

Reliability: Median test-retest reliabilities, based on a one-month interval, were reported as .89 for the total test and .64 to .88 for the subtests.

Validity: Correlations between PIAT total scores and the PPVT were reported between .53 and .79.

Publisher: American Guidance Service
Publisher's Building
Circle Pines, Minnesota 55014

Title: Peabody Picture Vocabulary Test

Date: 1965

Author: Lloyd M. Dunn

Range: Ages 2½-18

Administration: Individual; 10-15 minutes; 2 forms

Description: The PPVT was designed to give an indication of verbal intelligence as measured by receptive language vocabulary (listening vocabulary). It is composed of a graded series of 150 plates, each with 4 pictures. The examiner pronounces one stimulus word, and the subject indicates which of the four pictures best represents the stimulus word. Scores may be converted to IQ, mental age and percentile equivalents. According to Gearheart and Willenberg (1974), the PPVT has questionable appropriateness in use with some minority ethnic groups to which the stimulus pictures are not common. Hammill and Bartel state that although the PPVT yields and "IQ" score, it would be more proper to consider it a test of receptive vocabulary of standard English word meanings.

Development: The PPVT was standardized on 4,000 subjects of varying levels of intelligence over the age range of two years, six months through 18 years. Age norms were extrapolated downward to one year, nine months.

Reliability: Alternate-form reliability coefficients on the standardization population reportedly ranged from .67 at the 6-year level to .84 at the 17- and 18-year levels with a median of .77.

Validity: The median correlation of the PPVT with the Stanford-Binet was reported as .71; a correlation of .61 with the WISC was reported. The median correlation between the PPVT and school achievement was in the .50's.

Publisher: American Guidance Service, Inc.
Publisher's Building
Circle Pines, Minnesota 55014

Title: Perceptual Forms Test

Date: 1969

Author: Winter Haven Lions Research Foundation, Inc.

Range: Ages 5-8

Administration: Group; 10 minutes

Description: The PFT was originally published in 1955; thus, it is one of the earliest attempts to assess visual perception. It is one component of a perceptual and readiness evaluation and training program which utilizes template procedures. There are two parts in the PFT itself and Incomplete Forms. There are several versions of the PFT: group versions for kindergarten and grade one, an individual test for "beginning school children" and two versions for home use. Several handbooks have been written by different authors. It is questionable as to which, if any, is preferred or recommended. Little attention to accepted test construction procedures was apprently given by the developers of this approach. Both Mann (7th MMY) and Robinson (7th MMY) pointed out many deficiencies and confusing aspects of the procedures, handbooks, etc. Mann recommended that the authors start anew and provide norms, reliability figures and justifiable scoring guides in one explicit manual.

Development: No information was reported.

Reliability: No information was reported.

Validity: No information was reported.

Publisher: Winter Haven Lions Research Foundation, Inc.
P.O. Box 111
Winter Haven, Florida 33880

Title: Predictive Screening Test of Articulation

Date: 1975

Authors: Charles Van Riper and Robert L. Erickson

Range: Grade 1

Administration: Individual; 8 minutes

Description: The PSTA was devised to help the speech clinician identify those children who will overcome their articulatory difficulties by grade three without professional assistance.

Development: The initial group used in the development of this scale consisted of 167 beginning first grade children in Southwestern Michigan. All were judged, at that time, to have functionally defective articulation.

Reliability: A split-half reliability coefficient of .81 was reported for 293 cross-validation subjects.

Validity: Little research has been done on the predictive validity of the PSTA. The authors of the test emphasize that more research is needed.

Publisher: Western Michigan University
Continuing Education Office
Kalamazoo, Michigan

Title: Pre-Reading Screening Procedures

Date: 1969

Author: Beth H. Slingerland

Range: Grades K-1

Administration: Group; 40 minutes

Description: The purpose of the PSP is to identify children who may show indications of a "Specific Language Disability." The test consists of seven subtests which evaluate auditory, visual and visual-motor abilities. The manual recommends using the PSP with the *Metropolitan Readiness Tests* and the *Pintner-Cunningham Primary Test*. Jamison (7th MMY) recommended using one of the many readiness or intelligence tests available for pre-readers - one with more complete information. Kress (7th MMY) stated that the PSP merits consideration if much more data are gathered through other tests and informal information.

Development: No information was reported in the manual.

Reliability: No information was reported in the manual.

Validity: No information was reported in the manual.

Publisher: Educator's Publishing Service, Inc.
75 Moulton Street
Cambridge, Massachusetts 02138

Title: Pre-School Kindergarten Readiness Inventory

Date: 1971

Author: Margaret E. Green

Range: Grade K

Administration: Individual; 10 minutes

Description: The PKRI (also called the "Chula Vista") was developed for the purpose of measuring abilities and skills regarded as important for school success. It is preferably administered by the classroom teacher. Examples of skills surveyed are social and numerical awareness, motor performance, direction awareness, letter recognition, and body image awareness. The PKRI is easy to administer; a few easily located materials are required. The manual gives suggestions for interpretation of results of fall or spring assessment. In the revised edition, certain items are weighted to produce a possible score of 100.

Development: The PKRI was developed in the Chula Vista City School District, Chula Vista, California. The district will furnish score analyses of approximately 10,000 individually administered tests upon request. The PKRI was administered to approximately 2,000 students in 23 Chula Vista Schools in May, 1969. Approximately 20% were Spanish, 75% other Caucasian, and 5% Black, Oriental, and other non-Caucasian.

Reliability: No information reported.

Validity: A correlation of .83 was reported with the Metropolitan Readiness Tests, .55 with the Cognitive Abilities Test and .71 with the Cooperative Reading Test.

Publisher: Chula Vista City School District
84 East "J" Street
Chula Vista, California 92012

Title: Preschool Language Scale

Date: 1969

Authors: Irla Lee Zimmerman, Violette G. Steiner and Roberta L. Evatt

Range: Ages 1½-7

Administration: Individual; 30 minutes

Description: The PLS is designed to evaluate developmental progress, maturational language, strengths and deficiencies in the language skills of young children. It consists of a series of auditory and verbal language tasks at various age levels. Both Stark (7th MMY) and Ammons (7th MMY) suggested avoiding the PLS as there are better measures more carefully designed and with clear empirical evidence.

Development: No clear information was given regarding the development of the PLS.

Reliability: No information was reported.

Validity: No information was reported.

Publisher: Charles E. Merrill Publishing Company
1300 Alum Creek Drive
Columbus, Ohio 43216

Title: Primary Visual Motor Test

Date: 1970

Author: Mary R. Haworth

Range: Ages 4-8

Administration: Individual; 10-15 minutes

Description: The PVMT was designed to assess visual-motor development in the preschool and early primary grades. It consists of 16 stimulus cards which the subject is asked to reproduce. The PVMT is intended to be a downward extension of Bender-Gestalt concepts and method. According to Barclay (7th MMY) and Harris (7th MMY)-the PVMT appears to have merit for the assessment of visual-motor functions in young children; however, more data need to be gathered.

Development: The standardization population consisted of 100 children at each year from ages four through eight. The children were selected so that parental cocupations were representative of the 1960 U.S. census figures.

Reliability: Interscorer reliabilities were reported between .82 and .98 with an overall coefficient for all ages of .97. Test-retest reliability with a mean test interval of 52 days was .82.

Validity: Some evidence of validity is reported with small samples. Empirical evidence of predictive validity is limited. According to Harris, the main source of validity rests on construct and "face" validity.

Publisher: Grune and Stratton, Inc.
111 Fifth Avenue
New York, New York 10003

Title: Pupil Rating Scale

Date: 1971

Author: Helmer R. Myklebust

Range: Grades 3-4

Administration: Teacher-completed rating scale; completion time not reported

Description: The purpose of the Pupil Rating Scale is to screen learning disabled children. A score of 3 is average on the 5-point scale. The five areas evaluated are Auditory Comprehension, Spoken Language, Orientation, Motor Coordination, and Personal-Social Behavior. The author suggests that the scores should not be used for diagnostic purposes. Low scores may suggest further evaluation.

Development: This scale resulted from a 5-year Northwestern University research project which tested several identification procedures for selecting children with learning disabilities. Teachers' ratings were obtained for 2,000 third and fourth graders in four unidentified large suburban school systems. The age range of the sample was 7 through 10 years, with the majority being 8 and 9.

Reliability: No information was reported.

Validity: According to the manual, close agreement was found between PRS results and results from "intensive diagnostic evluations." The names of the diagnostic instruments were not mentioned and no data were given.

Publisher: Grune and Stratton, Inc.
111 Fifth Avenue
New York, New York 10003

Title: Purdue Perceptual Motor Survey

Date: 1966

Authors: Eugene G. Roach and Newell C. Kephart

Range: Ages 6-10

Administration: Individual; 20 minutes

Description: The PPMS was designed to assess a child's ability in jumping, identification of body parts, stepping stones, chalkboard tasks and other activities. The purpose of the test is to enable the teacher to identifiy children who are lacking perceputal-motor abilities. According to Hammill and Bartel (1975), the PPMS was probably never intended for use as a standardized instrument and would be better used as a structured informal device. The reviews in the Seventh Mental Measurement Yearbook are conflicting. The interested educator should consult this source for evaluative information.

Development: Two hundred children, grades one through four, who had no known motor defects and had not been referred for achievement evaluation comprised the normative sample. These children were all from the same school system in Indiana. Information on SES and sex is included in the manual.

Reliability: A test-retest reliability coefficient of .95 for 30 subjects was reported. Jamison (7th MMY) and Hammill and Bartel (1975) feel this may be spuriously high due to the small sample or extensive examiner training.

Validity: The concurrent validity coefficient for total score on Kephart's original scale and teacher ratings was reportedly .65.

Publisher: Charles E. Merrill Publishing Company
1300 Alum Creek Drive
Columbus, Ohio 43216

Title: Riley Articulation and Language Test

Date: 1971

Author: Glyndon D. Riley

Range: Grades K-2

Administration: Individual; 3 minutes

Description: The RALT was developed for the purpose of providing a fast screening device to determine which children most need speech therapy. It is not as comprehensive as other longer measures.

Development: The standardization population included 473 boys and girls from kindergarten through grade two. The subjects were low to middle socioeconomic level. Age, geographic and ethnic data are unspecified.

Reliability: A test-retest reliability coefficient of .81 for a one-week interval was reported for 76 children.

Validity: The RALT correlated .75 with the Templin-Darley Articulation Screening Test.

Publisher: Western Psychological Services
12031 Wilshire Boulevard
Los Angeles, California 90025

Title: Screening Tests for Identifying Children with Specific Language Disability

Date: 1970

Author: Beth H. Slingerland

Range: Grades 1-6

Administration: Group; 1 hour in 2 or 3 sessions; 4 levels (A, B, C, D)

Description: The Slingerland tests are designed to screen children with "specific language disabilities who are in need of special attention or remediation." According to Wepman (7th MMY), emphasis is placed on perceptual processes in the auditory, visual and kinesthetic modalities rather than on conceptual processes such as oral or written interpretation, vocabulary, or language patterns and usage. Deno (7th MMY) suggested that more empirical evidence is needed before the tests should be used for prediction or treatment.

Development: No information was reported in the manual.

Reliability: No information was reported in the manual.

Validity: No information was reported in the manual.

Publisher: Educator's Publishing Service, Inc.
75 Moulton Street
Cambridge, Massachusetts 02138

Title: Screening Test of Academic Readiness

Date: 1966

Author: A. Edward Ahr

Range: Ages 4 to 6 years, 5 months

Administration: Group; 50-60 minutes

Description: STAR was specifically designed to assess strengths and weaknesses related to school readiness in preschool and beginning kindergarten children. Language and performance activities are included. According to Huebner (7th MMY) it would be best used as an initial screening device if supplemented by other detecting and predictive measures. Magoon and Cox (7th MMY) feel that there would be little difference between the STAR and an IQ measure.

Development: The norms were developed using 1,500 preschool and kindergarten children between four years and six years, five months from a suburban white population of middle- to upper-class families.

Reliability: Test-retest reliabilities of .87 to .91 for four- to eight-week intervals were reported. Small samples were used.

Validity: Correlations of .72 and .67 (*Stanford-Binet*) and .76 (*Metropolitan Readiness Tests*) were reported.

Publisher: Priority Innovations, Inc.
P.O. Box 792
Skokie, Illinois 60076

Title: Silent Reading Diagnostic Tests

Date: 1970

Authors: Guy L. Bond, Bruce Balow and Cyril J. Hoyt

Range: Grades 2-6

Administration: Group; 90 minutes in 2 or 3 sessions

Description: These tests were designed to help the classroom teacher analyze specific silent reading abilities. The eight tests included are: Words in Isolation, Words in Context, Visual-Structural Analysis, Syllabication, Word Synthesis, Beginning Sounds, Ending Sounds and Vowel and Consonant Sounds. According to Bryant (7th MMY), these tests would probably be supplemented by individual tests. The tests provide useful, though limited, diagnostic information.

Development: The tests were standardized on 2,500 pupils described as representative of a population of approximately 38,000 pupils. The classes were from 10 cities in 3 states.

Reliability: Reliability coefficients of .80 to .95 were reported for the subtests. Only four classrooms were used to compute these coefficients.

Validity: No data for validity were reported.

Publisher:
Lyons and Carnahan
Rand McNally
P.O. Box 7600
Chicago, Illinois 60680

Title: Slosson Intelligence Test

Date: 1963

Author: Richard L. Slosson

Range: Ages 2 weeks and over

Administration: Individual; 10-20 minutes

Description: The SIT is designed to be a brief screening device of mental ability. Items are based on the Stanford-Binet Intelligence Scale and Gesell Developmental Schedules. The chronological age for calculating the IQ does not exceed 16 years, thus limiting the test for use with adults. There is a heavy emphasis on language skills, thus making the SIT more difficult for children who have language problems for cultural or individual reasons. This is especially true at the 2- to 4-year level. According to Hunt (7th MMY), one should not rely on the SIT in situations where important diagnostic decisions are required.

Development: The manual contains little information about the construction of the SIT. The standardization sample was not clearly described. The subjects were from New York; the sample size was not given in the description of the sample.

Reliability: A test-retest reliability coefficient of .97 within a two-month interval was reported.

Validity: According to Himelstein (7th MMY), only one true validity study was reported, showing correlations between the SIT and Stanford-Binet ranging from .90 to .98. These may be spuriously high since the SIT items for the ages reported (4-18) are Stanford-Binet adaptations.

Publisher: Slosson Educational Publications
140 Pine Street
East Aurora, New York 14052

Title: Specific Language Disability Test

Date: 1967

Author: Neva Malcomesius

Range: Grades 6-8

Administration: Group; 1 hour in 1 or 2 sessions

Description: This test is basically an upward extension of the *Screening Tests for Identifying Children with Specific Language Disabilities* by Slingerland. The reader is referred to the summary of the *Slingerland* tests as the information on description, development, reliability and validity are essentially the same.

Publisher: Educator's Publishing Service, Inc.
75 Moulton Street
Cambridge, Massachusetts 02138

Title: SRA Primary Mental Abilities Test

Date: 1962

Authors: L.L. Thurstone and Thelma Gwinn Thurstone

Range: Grades K-adult; 6 levels

Administration: Group; 1 hour and 15 minutes

Description: The SRA PMA was designed to provide "multifactored as well as general intelligence indices." The factors included are: Verbal Meaning, Number Facility, Reasoning, Perceptual Speed and Spatial Relations. The battery for adults is identical to the battery for grades 9-12 except for the title. Only one battery (grades 4-6) includes all five factors. According to Quereshi (7th MMY) and Schutz (7th MMY), the PMA is currently analogous to the Ford Model A. It is a "classical" battery but has failed to keep up with competing instruments. Both reviewers recommend other tests which are superior in terms of technical quality and functional utility.

Development: The standardization sample consisted of 32,000 children, ages 4-20 from public schools in five regions of the U.S. Data regarding sex and ethnic groups were omitted.

Reliability: Test-retest coefficients based on administrations of one-week and four-week intervals were reported. The coefficients for the total scores range between .83 and .95. The coefficients for individual factor scores vary considerably from grade to grade. Some fluctuations may be due to limited sample sizes (14 to 34). Perceptual Speed and Spatial Relations scores are more unstable than other scores.

Validity: According to Quereshi (7th MMY), there is little evidence presented in the technical report for the validity of PMA scores.

Publisher:

Science Research Associates
259 East Erie
Chicago, Illinois 60611

Title: Stanford-Binet Intelligence Scale, Third Revision

Date: 1964

Authors: Lewis M. Terman and Maud A. Merrill

Range: Ages 2 and over

Administration: Individual; 30-60 minutes; trained examiner

Description: The S-B is considered a measure of global or general intelligence. The subject proceeds through successive age levels as long as he is able to succeed in at least one task at a given level. The test provides a measure of the mental age and an intelligence quotient. The Wechsler Intelligence Scale for Children has seemingly surpassed the S-B in usage with older children. At present, no major, well standardized, general intelligence test of the point-scale type is available for children between 30 months, the ceiling of the Bagley Scales, and 48 months, the lowest age of the WPPSI (Freides, 7th MMY). According to Freides, it is here that the S-B serves a useful purpose today.

Development: In 1972, the S-B was administered to a stratified national sample of subjects from the ages of 2 years through 18 years. According to the publisher, the new norms tables resulted from the impact of recent social and cultural developments on test performance. These norms are available in a separate monograph.

Reliability: A discussion on reliability is included in the revised manual (1973).

Validity: A disucssion on validity is included in the manual.

Publisher: Houghton Mifflin Company
1900 South Batavia Avenue
Geneva, Illinois 60134

Title: Wechsler Intelligence Scale for Children-Revised

Date: 1974

Author: David Wechsler

Range: Ages 6 years to 16 years, 11 months

Administration: Individual; 1 hour; trained examiner

Description: The WISC-R is one of three *Wechsler Scales* based on the assumption that intelligence is global. These scales are designed to assess intelligence at various age levels. The 1949 WISC was revised due to changing population, changing environmental influences, item ambiguity, and various recommendations from those who had used the test for a period of years. The revision was also due in part to specific issues raised concerning test questions which might be unfair to minority populations. Approximately one-third of the test items are new or represent substantial modifications. The WISC-R continues to include verbal and performance tests and results in Verbal, Performance, and Full Scale IQ scores. The WISC-R cannot be administered with the 1949 WISC materials. According to Freides (7th MMY), the WISC is the best of the individual intelligence tests with special appeal in the separate verbal and performance scales. Osborne (7th MMY) stated that the WISC is the best intelligence test for children six to thirteen. He also stated that it is a major accomplishment that the WISC has survived the indictments against intelligence tests in general. According to Osborne the WISC is a "stable, general purpose individual intelligence test and is a useful and valid measure of immediate or present mental functioning." It remains to be seen if the WISC-R will prove to be more useful and valuable than the WISC.

Development: The standardization population of the WISC-R consisted of a stratified sample of children ages 6 years, 6 months through 16 years, 6 months. The sample included whites, as well as blacks and other nonwhite groups. The 1970 census data were used to specify a sample of 2,200 for the six stratification variables. These six variables were age, sex, race, occupation of head of household, urban-rural residence and geographic region.

Reliability: Split-half and test-retest coefficients were reported in the manual.

Validity: A correlation coefficient of .82 for 50 subjects was reported between the WISC-R and the WPPSI Full Scale IQs. A correlation coefficient of .95 for 40 subjects was reported between the WISC-R Full Scale and the WAIS Full Scale IQS. The average correlations for four ages between the WISC-R

Full Scale and the Stanford-Binet was reported as .73. The correlations in studies between the WISC-R and Stanford-Binet have been similar to values obtained in studies involving the 1949 WISC and Stanford-Binet.

Publisher:

Psychological Corporation
304 East 45th Street
New York, New York 10017

Title: Stanford Early School Achievement (Level I)

Date: 1969

Authors: Richard Madden and Eric F. Gardner

Range: Grades K-1

Administration: Group; 1 hour and 30 minutes; 3-5 sessions

Description: The SESAT I was designed to provide a measure of the child's cognitive abilities. According to the manual, the SESAT I is not a readiness test. It consists of four subtests: The Environment, Mathematics, Letter and Sounds and Aural Comprehension. According to Hagen (7th MMY) and Mehrens (7th MMY), the SESAT I should be useful in determining placement and instruction.

Development: The final norm sample consisted of 8,300 kindergarten pupils and 11,100 first graders in 27 states. The school systems were selected to be representative in location, size and socioeconomic level.

Reliability: Split-half reliabilities were reported ranging from .76 to .85.

Validity: No information was reported.

Publisher: Harcourt Brace Jovanovich
Test Department
757 Third Avenue
New York, New York 10017

Title: Stanford Diagnostic Arithmetic Test

Date: 1966

Authors: Leslie S. Beatty, Richard Madden, and Eric F. Gardner

Range: Grades 2.5-8.5; 2 levels

Administration: Group; 2 hours, 50 minutes in 6 sessions (Level I), 3 hours, 50 minutes in 7 sessions (Level II)

Description: The SDAT was designed to identify and diagnose specific weaknesses in arithmetic. The focus is on an understanding of properties of the number system and on computation. It is intended to be given early in the school year. According to Rogers (7th MMY), there is little research on the effectiveness of group diagnostic arithmeitc tests. He cautions users concerning the use and interpretation of the SDAT.

Development: The standardization population consisted of 8,000 pupils in four school systems. The extent to which this sample represents the national school population is not defined.

Reliability: Split-half reliability coefficients were reported at each grade for all subtests. The reliabilities ranged from .89 to .98 for total scores. Several subtest reliabilities were in the .70's while two were .57 and .60.

Validity: Most of the correlations between the SDAT and the Stanford-Achievement Tests are high for measuring achievement. No evidence of diagnostic validity is provided.

Publisher: Harcourt Brace Jovanovich
Test Department
757 Third Avenue
New York, New York 10017

Title: Stanford Diagnostic Reading Test

Date: 1968

Authors: Bjorn Karlsen, Richard Madden and Eric F. Gardner

Range: Grades 2.5-8.5; 2 levels

Administration: Group; 2 hours and 20 minutes in 4 sessions (Level I);
1 hour and 30 minutes in 3 sessions (Level II);
2 forms for each level

Description: The SDRT is composed of seven subtests on the Level I test and eight on Level II. Both levels contain comprehension, vocabulary, sound discrimination, syllabication and blending. Level I measures general comprehension only, while Level II measures literal and inferential comprehension. The manual advises teachers against using individual items for diagnosis.

Development: An item analysis was conducted on 15,000 pupils in five states. The standardization population consisted of 12,000 pupils from six school systems. Very little other information was given.

Reliability: Split-half reliabilities were reported ranging from .87 for grade 5 to .94 for grade 3.

Validity: Validity of the SDRT was discussed; however, almost no data were reported.

Publisher: Harcourt Brace Jovanovich
Test Department
757 Third Avenue
New York, New York 10017

Title: Standard Reading Inventory

Date: 1966

Author: Robert A. McCracken

Range: Grades 1-7

Administration: Individual; 30 minutes; 2 alternate forms

Description: This inventory consists of eleven stories for oral reading, eight stories for silent reading, and eleven word lists for measuring sight vocabulary. The test is designed to measure word recognition in isolation and context, errors in oral reading, comprehension, word meaning and speed of oral and silent reading. According to H.A. Robinson (7th MMY), the SRI would be useful as a rough, semidiagnostic tool. It would provide more information about the reading process than would a group standardized silent reading test.

Development: This test is not standardized and there was no information about specific norming procedures.

Reliability: Alternate-form reliabilities ranged from .86 to .91 for the level scores and .68 to .99 (median .93) for the subtest scores.

Validity: A discussion of content validity is included in the manual. Content validity was obtained through a) vocabulary control; b) basing sentence length, content, and general style on three basal reading series and c) the use of Spache and Dale-Chall readability formulas.

Publisher: Klamath Printing Company
320 Lowell Street
Klamath Falls, Oregon 97601

Title: Templin-Darley Tests of Articulation

Date: 1969

Authors: Mildred C. Templin and Frederic L. Darley

Range: Ages 3 and over

Administration: Individual; 5 minutes for screening test; 20 minutes for total diagnostic test

Description: The TDTA may be used as a screening device (50 items) or as a diagnostic instrument (141 items). According to Haller (7th MMY), the 1969 TDTA is probably " . . . the best published clinical measure of phoneme acquisition in terms of its rationale, the variety and quality of normative data, and flexibility. Its apparent limitations are common to most articulation tests."

Development: The norms in the 1969 edition were based on data gathered in 1957 for the first edition. The items were administered to 480 children ages 3 through 8. The subjects were white, average intelligence, with no gross hearing loss, and were enrolled in 14 public schools and 21 nursery schools in Minneapolis and St. Paul. The subjects were selected to be representative of the U.S. urban population.

Reliability: Test-retest reliability coefficents of .93 to .99 on single age groups between 2 and 5 years old for a 8-day interval were reported.

Validity: The manual contains some information on validity.

Publisher: Bureau of Educational Research and Service
Division of Extension and University Services
University of Iowa
Iowa City, Iowa 52240

Title: Utah Test of Language Development

Date: 1969

Authors: Merlin J. Mecham, J. Lorin Jex and J. Dean Jones

Range: Ages 1 year, 5 months to 14 years, 5 months

Administration: Individual; 30-45 minutes

Description: The UTLD is a test of general level of language functioning in normal or handicapped children. It tests expressive and receptive language skills. It also contains items that test conceptual development. The test should probably not be given to children with visual-perception problems, inner-city children, or children of minority racial or ethnic groups. According to Butler (7th MMY), the UTLD is a useful screening device of language skills, especially at the preschool level.

Development: The standardization population consisted of 273 normal white subjects, ages 1 year, 6 months to 14 years, 5 months. All socioeconomic levels were represented.

Reliability: A split-half reliability coefficient of .93 was reported.

Validity: According to the authors, test items have "face" validity as they were selected from previously standarized sources.

Publisher: Communication Research Associates, Inc.
Salt Lake City, Utah

Title: Valett Developmental Survey of Basic Learning Abilities

Date: 1966

Author: Robert E. Valett

Range: Ages 2-7

Administration: Individual; 60 minutes

Description: The Valett Survey consists largely of a selection and adaptation of items from other scales such as those by Gesell, Kephart, Binet, Frostig and Wechsler. There are 233 tasks covering seven basic areas: motor integration and physical development, tactile discrimination, auditory discrimination, visual-motor coordination, visual discrimination, language development and verbal fluency, and conceptual development. According to the manual, the survey results are to be used to determine if further diagnostic evaluation is needed and to plan an educational program for the child. Ruth (7th MMY) stated that most users would have difficulty scoring the items, as directions in the manual are inadequate. Mann (7th MMY) and Ruth both comment on the frequent violations of educational test standards in the development of this device.

Development: No clear rationale was presented for the selection or age-level assignment of items. No information was offered on normative sample or the appropriateness of item placement.

Reliability: No information was reported.

Validity: No information was reported.

Publisher: Consulting Psychologists Press, Inc.
577 College Avenue
Palo Alto, California 94306

Title: Vane Kindergarten Test

Date: 1968

Author: Julia R. Vane

Range: Ages 4 years to 6 years, 11 months

Administration: Two parts may be group administered; one part must be administered individually; 30 minutes; trained examiner

Description: The purpose of the VKT is to "evaluate the intellectual and academic potential and behavior adjustment of young children." The three parts included are a perceptual motor subtest, a draw-a-man subtest and a vocabulary subtest. The vocabulary subtest is administered individually. All three subtests contain items similar to those widely used in tests for young children.

Development: The norms are based on 400 subjects, all from New York and New Jersey. The sample is representative of rural-urban, white-nonwhite, and occupational groups.

Reliability: Test-retest reliability coefficients were .97 after one week and .88 after five months, for small, undefined samples.

Validity: The VKT correlated .76 with the Stanford-Binet given to 212 subjects with the tests given two years apart.

Publisher: Clinical Psychologists Publishing Company, Inc.
4 Conant Square
Brandon, Vermont 05733

Title: Vineland Social Maturity Scale

Date: 1965

Author: Edgar A. Doll

Range: Birth to adult

Administration: Individual; 20-30 minutes

Description: The Vineland is an informant-interview technique used for the screening and diagnosis of degree or level of social competence. Items are arranged in order of increasing average difficulty in six areas: self-help (general, eating and dressing), self-direction, occupation, communication, locomotion, and socialization. The parent, teacher, or counselor scores each of 117 items. Although a Social-Age Value up to 30 years may be obtained, the scale is more frequently used with young children. According to Cruickshank (4th MMY) and Teagarden (4th MMY), the Vineland can add to clinical insights regarding an individual.

Development: The standardization sample was from the eastern U.S. The norms are based on 20 normal subjects at each of 31 age levels, a total of 620 subjects.

Reliability: Test-retest coefficients reportedly ranged between .94 and .99 based on 1.35 years between tests.

Validity: According to the publisher's information, if the range of informants about the child's behavior is increased from parents to educational and psychological personnel, agreement of evaluation correlated .92.

Publisher: American Guidance Service
Publisher's Building
Circle Pines, Minnesota 55014

Title: Wechsler Preschool and Primary Scale of Intelligence

Date: 1967

Author: David Wechsler

Range: Ages 4-6½

Administration: Individual; 50-60 minutes; trained examiner

Description: The WPPSI is, in a sense, a downward extension of the WISC and is also designed to assess a child's global intelligence. It is similar in many ways to the other *Wechsler Scales*. It results in Verbal, Performance and Full Scale IQ scores. It contains five verbal and five performance subtests. Some of the subtests and items have been adapted from the WISC for use with younger children. According to Eichorn (7th MMY), "the WPPSI is the best standardized and most up-to-date individual test available." However, due to an inadequate floor (at the lower end of the scale) it is not appropriate to use for differentiating among the moderate and severely retarded.

Development: The standardization population consisted of 600 boys and 600 girls stratified with respect to geographic region, urban-rural residence, and father's occupation based on the 1960 census. Both white and nonwhite subjects were included.

Reliability: Test-retest reliability coefficients were reported for the Verbal (.86), Performance (.89) and Full Scale IQ (.92) from a sample of 50 children retested after an average of eleven weeks. Split-half coefficients were reported at each age level for every subtest except Animal House. The average split-half reliability coefficients were .94 (Verbal), .93 (Performance) and .96 (Full Scale).

Validity: A correlation coefficient of .76 with the *Stanford-Binet* was reported for 100 five- to six-year olds. The Verbal IQ correlated more highly with the *Stanford-Binet* than did the Performance IQ. Correlations were also reported with the PPVT and the *Pictorial Test of Intelligence*.

Publisher:
Psychological Corporation
304 East 45th Street
New York, New York 10017

Title: Wide Range Achievement Test

Date: 1965

Authors: J.F. Jastak, S.R. Jastak and S.W. Bijou

Range: Ages 5-adult; 2 levels

Administration: Individual; 20-30 minutes

Description: The WRAT was designed as a method to assess reading (word recognition and pronunciation), written spelling and arithmetic computation. Merwin (7th MMY) felt that the WRAT is potentially useful in a clinical setting but is impractical for general school use. According to Thorndike (7th MMY), the test might be valuable in a clinical or research setting but he would hesitate to recommend it for other purposes.

Development: A group of 5,800 children and adults from seven states comprised the standardization population. The authors stated that "no attempt was made to obtain a representative national sampling."

Reliability: Split-half reliability coefficients of .98-.99 were reported. Reviewers Merwin and Thorndike question these figures as certain features of the WRAT design tend to inflate split-half reliabilities. Correlations between Level 1 and Level 2 were reported in the range of .85 to .90/

Validity: The manual has a section concerning validity. However, Thorndike stated that it is "hard to reconcile these statements with each other or with the usual concepts of test validation." Merwin is also highly critical of the narrative on validity.

Publisher: Harcourt Brace Jovanovich
7555 Caldwell
Chicago, Illinois 60648

Title: What I Like to Do

Date: 1954

Authors: Marcella R. Bonsall, Charles E. Meyers and Louis P. Thorpe

Range: Grades 4-7

Administration: Individual inventory; 40 minutes

Description: *What I Like to Do* is an interest inventory which yields an interest pattern in eight areas of in-school and out-of-school activities. These areas are: art, music, social studies, active play, quiet play, manual arts, home arts and science. The purpose of the inventory is to provide a means of identifying pupil interests so they may be used in guidance and instruction.

Development: The standardization population was stratified for grade, sex, urban-rural, SES and geographic region. Thirty-eight hundred children from 33 states and 51 schools were administered the inventory.

Reliability: Coefficients of reliability for the eight areas for boys and girls in each grade ranged from .70 to .97.

Validity: No data were reported on validity. The manual includes a discussion of the item selection procedures.

Publisher: Science Research Associates
259 East Erie
Chicago, Illinois 60611

Title: Woodcock Reading Mastery Tests

Date: 1973

Author: Richard W. Woodcock

Range: Grades K-12

Administration: Individual; 20-30 minutes; two forms

Description: The WRMT is a battery of five criterion-referenced reading tests. The tests are letter identification, word identification, word attack, word comprehension, and passage comprehension. The emphasis of test interpretation is on predicted individual performance, rather than comparison with others.

Development: Item analysis was based on 35,000 individually administered tests to subjects in kindergarten through grade 12 in Illinois, Michigan, Minnesota and Tennessee. Normative data were obtained on 5,000 subjects selected from a stratified random sampling plan from across the U.S.

Reliability: Split-half reliabilities were computed for 850 pupils in grades one, four and ten. Reliabilities for the five tests fell between .90 and .99. Total test score reliabilities fell in the .97 to .99 range. Test-retest reliabilities with alternate forms ranged from .68 to .97. The tests were given to 200 pupils in grades two and seven with one week intervening.

Publisher: American Guidance Service, Inc.
Publisher's Building
Circle Pines, Minnesota 55014